

1 / 110

FIG. 1

ID-65

Clone 3-60

GTGTTTATGATGAAAAAAGGACAAGTAAATGATACTAAGCAA
TCTTACTCTCTACGTAAATATAAATTTGGTTTAGCATCAGTAA
TTTTAGGGTCATTCATAATGGTCACAAGTCCTGTTTTGCGGA
TCAAACCTACATCGGTTCAAGTTAATAATCAGACAGGCACTAG
TGTGGATGCTAATAATTCTTCCAATGAGACAAGTGCGTCAAGT
GTGATTACTTCCAATAATGATAGTGTTCAAGCGTCTGATAAAG
TTGTAAATAGTCAAAATACGGCAACAAAGGACATTACTACTC
CTTTAGTAGAGACAAAGCCAATGGTGGAACAAACATTACCTG
AACAAAGGGAATTATGTTTATAGCAAAGAAACCGAGGTGAAAA
ATACACCTTCAAAATCAGCCCCAGTAGCTTTCTATGCAAAGAA
AGGTGATAAAGTTTTCTATGACCAAGTATTTAATAAAGATAAT
GTGAAATGGATTTTCATATAAGTCTTTTGGTGGCGTACGTCGAT
ACGCAGCTATTGAGTCACTAGATCCATCAGGAGGTTTCAGAGA
CTAAAGCACCTACTCCTGTAAACAAATTCAGGAAGCAATAATC
AAGAGAAAATAGCAACGCAAGGAAATTATACATTTTACATA
AAGTAGAAGTAAAAAATGAAGCTAAGGTAGCGAGTCCAATC
AATTTACATTGGACAAAGGAGACAGAATTTTTTACGACCAAA
TACTAACTATTGAAGGAAATCAGTGGTTATCTTATAAATCATT
CAATGGTGTTTCGTCGTTTTGTTTTGCTAGGTAAAGCATCTTCA
GTAGAAAAAACTGAAGATAAAGAAAAAGTGTCTCCTCAACCA
CAAGCCCGTATTACTAAAACCTGGTAGACTGACTATTTCTAACG
AAACAACCTACAGGTTTTGATATTTTAATTACGAATATTAAAGA
TGATAACGGTATCGCTGCTGTTAAGGTACCGGTTTGGACTGAA
CAAGGAGGGCAAGATGATATTAATGGTATACAGCTGTAACCT
ACTGGGGATGGCAACTACAAAGTAGCTGTATCATTGCTGAC
CATAAGAATGAGAAGGGTCTTTATAATATTCATTTATACTACC
AAGAAGCTAGTGGGACACTTGTAGGTGTAACAGGAACTAAAG
TGACAGTAGCTGGAACCTAATTCTTCTCAAGAACCTATTGAAAA
TGGTTTACCAAAGACTGGTGTTTATAATATTATCGGAAGTACT
GAAGTAAAAAATGAAGCTAAAATATCAAGTCAGACCCAATTT
ACTTTAGAAAAAAGGTGACAAAATAAATTATGATCAAGTATTG
ACAGCAGATGGTTACCAAGTGGATTTCTTACAAATCTTATAGTG
GTGTTTCGTCGCTATATTCCTGTGAAAAAGCTAACTACAAGTAG
TGAAAAAGCGAAAGATGAGGCGACTAAACCGACTAGTTATCC
CAACTTACCTAAAACAGGTACCTATACATTTACTAAAACCTGTA
GATGTGAAAAGTCAACCTAAAGTATCAAGTCCAGTGGAATTT
AATTTTCAAAAGGGTGAAAAAATACATTATGATCAAGTGTTA
GTAGTAGATGGTCATCAGTGGATTTTCATACAAGAGTTATTCCG
GTATTCGTCGCTATATTGAAATTTAA

FIG. 1 CONT'D

3 / 110

MILRRRTIVLWQLGIAISLILSILALNLYFHSTPLQTNAALRNLAPSLNHLFGTD
 GLGRDMFVRTIKGLYFSLQVLLGALMGVILATVFGVLAGLGNHDKIIAWL
 VDLFIGMPHLIFMILISFVVVGKGAQGVIIATAVTHWPSLARLRNEVYHLKNKE
 FVQLSKSMGKTPYYIVRHHILPLIASQIFIGFILLFPHVILHEASMTFLGFGLSAE
 QPSVGIIILSEAAKHISLGNWWLVIFPGLYLILVVNAFDTIGESLKKLFYPQTDHF
 *

Sequence description

- A) Length: 822 bp - 274 aa (full length gene)
- B) Sequence Characteristics:
 Potential leader peptide sequence
 Orf is preceded by a potential Shine-
 Dalgarno sequence.

ID-78

Clone 3-5b

ATGACAGAAACATTATTAAGCATTAAAGACCTCTCCATCACCTTCACTCAA
 TACGGAAGATTTTTAAAACCATTTCAATCAACACCGATACAAGCGCTGA
 ATTTAGAAATTAAAAAAGGTGAGTTATTAGCTATTATAGGTGCTAGTGGTT
 CGGGGAAGAGTTTATTAGCACATGCTATTATGGATATTCTTCCTAAAAATG
 CATCTGTAACAGGAGATATGATTTATCGTGGTCAATCACTAAATTCTAAAC
 GCATTAAACAGTTGCGAGGAAAAGATATTACGTTGATTCCACAATCAGTTA
 ATTATTTAGATCCATCTATGAAAGTCAAACATCAGGTGCGCTTAGGTATCT
 CAGAAAATTCAAAGGCTACTCAAGAAGGATTGTTTCAACAGTTTGGTTTAA
 AAGAAAGTGATGGTGACTTGGATCCTTTCCAACCTTTCTGGCGGAATGCTCC
 GACGTGTTTGTGTTACAACGTGTATTAGTGATAAGGTTTCTTTGATTATTGC
 GGATGAGCCCAACCCTGGATTACATCCAGATGCTCTGCAAATGGTTTTAGA
 CCAACTACGCTCCTTTGCAGATAAAGGAATAAGCGTTATATTTATCACTCA
 TGATATTGTAGCAGCTAGTCAAATTGCTGATCGTATTACTATTTTAAAGA
 GGGAAAAGCTATTGAAACAGCTCCAGCTAGTTTCTTTAGCGGAAATGGAG
 AGCAGTTACAAACAGAATTTGCTAGAAAGTTTATGGCGCTCTCTCCACAGC
 AAGAATTTTGAAGGAGTTACTCATGACCTTAGAGGCTAA

MTETLLSIKDLSITFTQYGRFLKPFQSTPIQALNLEIKKGELLAHIGASGSGKSL
 AHAIMDILPKNASVTGDMYRQSLNSKRIKQLRGKDITLIPQSVNYLDPSMK
 VKHQVRLGISENSKATQEGLFQQFGLKESDGDLDPFQLSGGMLRRVLFTTCIS
 DKVSLIIADEPTPGLHPDALQMVLDQLRSFADKGISVIFITHDIVAASQIADRITI
 FKEGKAJETAPASFFSGNGEQLQTEFARSLWRSPLQQEFLKGVTHDLRG*

FIG. 1 CONT'D

FIG. 1 CONT'D

FIG. 1 CONT'D

7 / 110

ID-82

Clone 48

ATGGCAGATAAAAAACAGAACATTTAACTTGTAGGTGCAGGATCTTCTAG
CACACAAGAAAAAATTGAAAAGCCTGCTCTTTCGTTTATGCAAGATGCGTG
GCGTCGCTTGAAAAAAAACAAATTAGCAGTAGTTTCACTCTATTTATTAGC
TCTTTTACTTACTTTTTTCGTTAGCCTCAAATTTATTTGTAACTCAGAAGGAT
GCTAATGGGTTTGATTGAAAAAAGTAACGACATATCGCAACTTACCACCT
AAATTGAGTTCAAACCTTCCTTTTTGGAATGGTAGCATTAAATCCATCA

MADKNRTFKLVGAGSSSTQEKIEKPALSFMQDAWRRLKKNKLAVVSLYLLA
LLLTFSLASNLFVTQKDANGFDSKKVTTYRNLPPKLSSNLPFWNGSINPS

Sequence description:

A) Current length is 303 bp - 101 aa

B) No obvious signal peptide but Shine

Dalgarno sequence upstream of the ATG start

codon. Not identified directly using the LEEP system but was found

directly downstream of ID-34 described in WO 00/06736.

ID-83

Clone 98

ATGAAAATAGTAGTACCAGTAATGCCTCGCAGTCTTGAAGAGGCTCAAGA
AATAGATTTATCAAAATTTGATAGTGTTGATATTATTGAATGGCGAGCTGA
TGCTTACCAAAGGATGACATTATTAATGTAGCTCCAGCTATTTTTGAGAA
ATTCGCAGGTCATGAAATTATTTTTACTTTTCGTACAACGCGTGAAGGTGG
TAATATTGTCTTATCTGATGCTGAGTATGTTGAGTTAATCCAGAAAATTAA
TTCTATCTACAATCCAGATTATATTGATTTTGAAGTATTTTTCACATAAAGAA
GTTTTTCAAGAAATGCTAGAATTTCCAAATTTAGTCCTGTCTTATCACAATT
TTCAAGAGACACCGGAGAATATTATGGAGATATTTTCAGAATTAACAGCC
CTAGCACCACGAGTTGTGAAAATCGCAGTAATGCCAAAGAATGAACAAGA

FIG. 1^{CONT'D}

8 / 110

TGTCTTAGACGTTATGAATTACACTCGCGGTTTCAAGACTATTAATCCTGA
TCAAGTTTATGCGACGGTATCTATGAGTAAAATTGGACGTATTTCTCGTTTT
GCTGGTGATGTAAGTGGATCTAGTTGGACATTTGCATATTTAGATTCATCT
ATCGCACCCGGACAAATTACTATTTTCAGAGATGAAGCGTGTCAAAGCATT
GCTTGACGCTGACTGA

MKIVVPVMPRSLEEAQEIDLSKFDSVDIIEWRADALPKDDIINVAPAIFEKFAG
HEIIFTFRTTREGGNIVLSDAEYVELIQKINSIYNPDYIDFEYFSHKEVFQEMLEF
PNLVLSYHNFQETPENIMEIFSELTAAPRVVKJAVMPKNEQDVLDVMNYTRG
FKTINPDQVYATVSMKIGRISRFAAGDVTGSSWTFAYLDSSIAPGQITISEMKRV
KALLDAD*

Sequence description:

- A] Length 678 bp, 225 aa (full length gene)
- B] No obvious signal peptide, but there is a Shine Dalgarno immediately upstream of ORF.

ID-84

Clone RS-52

ATGAAAGACTTATTTGCAACAACAGAAGCATCATCAAGGAAACAGGAACA
AGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAACAAATGGTA
ATCAAATAAAAAAAAAATTGAGTTTATCGACTTTCAAAAAAATGAGATGACA
GGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCGATTAGTTTT
TCTGAAGATAGAATTGGTGGTAACTTAGAGCATTAGGATATCAACCGAA
TGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCAAAATGTTAATGA
TATTGAAGTGATTTATATGAAGAAAGAATAG

MKDLFATTEASSRKQEQDRIVNYIKQHVELTNGNQIKKIEFIDFQKNEMTGTW
GISTKINEQFSISFSEDRIIGGKLRALGYQPNEIGFSKDINSNNQNVNDIEVIYMK
KE*

Sequence description:

- A] length: 333 bp - 111 aa (partial sequence)
- B] No obvious Shine Dalgarno sequence upstream of the ATG start codon, and no obvious signal peptide within the protein.

FIG. 1_{CONT'D}

9 / 110

ID-85

Clone RS-53

ATGAAAAAACGTATATGGTATTTGATAATAATAATCACAGTAATTTTAGGA
GGACTAGCCATGAAAAACTTATTTGCAACAACAGAAGCATCATCAAGGAA
ACAGGAACAAGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAA
CAAATGGTAATCAAATAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAT
GAGATGACAGGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCG
ATTAGTTTTTCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATAT
CAACCGAATGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCA

MKKRIWYLIHITVILGGLAMKNLFATTEASSRKQEQDRIVNYIKQHVELTNGN
QIKKIEFIDFQKNEMTGTWGISTKINEQFSISFSEDRI GGKLRALGYQPNEIGFSK
DINSNNQ

Sequence description:

- A) Length: 351 bp - 117 aa (Partial sequence)
- B) Obvious signal peptide and Shine Dalgarno sequence upstream of the ATG start codon.

ID-86

Clone ID-74

ATGTCAAATCAATATGATTATATCGTTATTGGTGGAGGTAGT
GCAGGCAGTGGTACCGCTAATAGGGCAGCCATGTATGGAGC
AAAAGTCCTGTTAATTGAAGGTGGACAAGTAGGTGGAACCTTG
TGTTAACTTAGGTTGTGTACCTAAGAAAATCATGTGGTATGG
TGCACAAGTTTCTGAGACACTCCATAAGTATAGTTTCAGGTTA
TGGTTTTGAAGCCAATAATCTTAGTTTGTATTTTACTACTCTA
AAAGCTAATCGCGATGCTTACGTGCAGCGGTCTAGACAGTCG
TATGCCGCTAATTTTGAGCGTAATGGGGTCGAAAAGATTGAT
GGATTTGCTCGTTTTATTGATAACCATACTATTGAAGTGAATG
GTCAGCAATATAAAGCTCCTCACATTACTATTGCAACAGGTG

FIG. 1_{CONT'D}

FIG. 1 CONT'D

FIG. 1 CONT'D

12 / 110

TTAAAAATAATCGAAGGTTTAGATTACTCTACATTACGTCAGACTACTCAA
 ACTCAGTTTAGAGAAATGACTACTGCAGGTAAAGCGTTGTTATCAAAATCT
 AACCTCCGACAAGGAGAAAAACAAAATTCTTCAAGAAGCACAATTTTT
 CCTTGGTCGCGTTGATTTGGATAAAGCCATAGCTAAAGCTGAGAAGGCTTT
 AGTGACCAAGAAGGCAACAAAGAATGGTCATTTGCTTGAGAGGAGTATTA
 ACAAAGCGGTATTAGCTTATAATAATAGTGCTATTAAGGCTAATGTGA
 AGCGCTTGGAAGGAGGTTAGACTTGCTGACAGATTTAGTCGAGGGAAAA
 GGACCATTAGCGCAAGCTACAATGGTACAAGGAGTTTATTTATTAAGAC
 GCCTTTACCATTGCCAGAATATTATATCGGATTGAACGTTTATTTTGACAA
 GTCTGGAAAATTGATTTATGCACCTTGATATGAGTGATACTATTGGCGAGGG
 AAAAAAGATGCATATGGTAATCCTATATTAATGTTGACGAGGATAATG
 AAGGTTATCATACCTTGGCAGTTGCCACTTTAGCTGATTATGAAGGTCTTT
 ATATTAAGATATTTTAAATAGTTCCTTGATAAGATTAAAGCAATACGCC
 AGATTCCTTTGGCAAAATATCATAGATTAGGAATTTCCAAGCTATCCGAA
 ATGCAGCGGCAGAAGCAGACCGATTGCTTCCTAAGACACCTAAGGGGTAC
 CTAATGAAGTCCCAAATTATCGTAAAAAACAAATGGAGAAAAATTTAAA
 ACCAGTTGATTATAAAACGCCGATTTTAATAAGGCTTTACCTAATGAAAA
 GGTAGACGGTGATAGAGCGGCTAAAGGTCATAATATAAATGCGGAGACTA
 ATAATTCTGTAGCTGTAACACCAATAAGGTCCGAGCAGCAATTACATAAGT
 CACAGTCTGATGTAAATTTACCTCAAACAAGTTCTAAAAATAATTTTATAT
 ACGAGATTCTAGGATACGTTAGTTTATGTTTGCTTTTCCTAGTAAGTCTGG
 GAAAAAGGAAAACGAGCAAGAAAAATAA

MTKKHLKTLALALTTVSVVTYSQEVYGLEREESVKQEQTQSASEDDWFEEDN
 ERKTNVSKENSTVDETVSDLFSDGNSNNSSSKTESVVSDDPKQVPAKPEVTQE
 ASNSSNDASKVEVPKQDTASKKETLETSTWEAKDFVTRGDTLVGFSKSGINKL
 SQTSHLVLP SHAADGTQLTQVASFATPDKKTAAEYTSRLGENGKPSRLDIDQ
 KEIIDEGEIFNAYQLTKLTIPNGYKYSIGQDAFVDNKNIAEVNLPESLETISDYAF
 AHMSLKQVKLPDNLK VIGELAFFDNQIGGKLYLPRHLIKLAERAFKSNRIQTV
 EFLGSKLK VIGEASFQDNNLRNVMLPDGLEKIESEAFGNPGDEHYNNQVVL R
 TRTGQNP HQ LATENTYVNPDKSLWRATPDMDYTKWLEEDFTYQKNSVTGFS
 NKGLQKVR RNKNLEIPKQHNGITITEIGDNAFRNVDFQSKTLRKYDLEEIKLPS
 TIRKIGAF AFQSN NLKSFEASEDLEEIKEGAFMNNRIGTLDLKD KLIKIGDAAFH
 INHIYAIVLPESVQEIGR SFRQNGALHLMFIGNKVKTIGEMAFLSNKLESVNL
 SEQKQLKTIEVQAFSDNALSEVVLPPNLQTIREEAFKR NHLKEVKGSSTLSQITF
 NAFDQNDGDKRFGKKVVVRTHNNSHMLADGERFIIDPKLSSTMVDLEKVL
 KIEGLDYSTLRQTTQTQFREMTTAGKALLSKSNLRQGEKQKFLQEAQFFLGR
 VDLDKAIKAEKALVTKKATKNHLLERSINKAVLAYNN SAIKKANVKRLEK
 ELDLLTDLVEGKGPLAQATMVQGVYLLKTPLPLPEYYIGLNVYFDKSGKLIYA
 LDMSDTIGEGQKDAYGNPILNVDEDN EGYHTLAVATLADYEGLYIKDILNSSL
 DKIKAIRQIPLAKYHRLGIFQAIRNAAAEADRLLPKTPKGYLNEVPNYRKKQM
 EKNLKPVDYK TPIFNKALPNEKVDGDRAAKGHNINAETNNSVAVTPIRSEQQL
 HKSQSDVNLPQTSSKNNFIYEILGYVSLCLLFLVTAGKKGKRARK*

FIG. 1 CONT'D

13 / 110

Sequence description:

- A] Length 3168 bp - 1056 aa (Partial sequence)
- B] Obvious signal peptide with Shine Dalgarno sequence upstream of the ATG start codon.

ID-88

Clone RS-56

GCAGGATACATCATGCACAAGCACGAGGCTATCGTGTCATGCTGGGGTCA
ACCCAGGAAGACATGTCGGCACAAGCTGAAGATTTCTTTACAGTCTGTACA
CAATAAAGAGACGGGTAAGAGCGCTTTTAATGACAAAGAACGACTAGCAA
TT

AGYIMHKHEAIVSCWGQPRKTCRHLKISLQSVHNKETGKSAFNDKERLAI

Sequence description:

- A] Length: 153 bp - 51 aa (partial sequence)
- B] No signal peptide visible, insufficient sequence data to determine the presence of a Shine Dalgarno sequence.

ID-89

Clone RS-58

GTGTCATTTATGCAAAGAAAATCCTATTTAAAATCCATGAGTGTTCTTACT
TTAACAGCTTGTCTTATATCAGGATATGTGGTTAAAGATATTGCTATGTTA
CATGCAGTATCTGCCAGTGAGAAGAAAGCAAATAATGTCAGTCCGAGAGA
AAATCTCTACAGGGCTGTCAATGATAATTGGCTAGCCAATACAAAACCTCA
AACAAAGGGCAGACTAGTGTTAATAGTTTTTCAGAAATTGAGGATAAATTA
AAGCAACTGTTAGTGTCTGATATGGCTAAAATGGCCTCAGGAAAGATTGA

FIG. 1_{CONT'D}

14 / 110

AACAACCAATGATGAACAGAAAAAATGGTTGCATACTATAAACAAGGTA
TGGACTTTAAAACAAGAGATAAAAAATGGTCTCAAACCTCTAAAACCAGTT
TTACAAAAACTTGAAGCAGTCTCTTCAATGAAAGACTTTCAAAGTTTGGCC
CATGATTTTGTGATGAGTGGTTTTGTTTTACCATTTGGTTTGACTGTGGAAA
CCAATGCTCGAGATAATAGCCAAAAGCAATTGGTGCTTCGTCAAGCACCC
GCATTACTTGAATCACCTGACCAATATAAGAAGGGCAATAAAGAAGGTGA
GGCTAAATTATCAGCTTACCGTACTTCAGCAATGGCTTTGCTTAAACAAGC
TGGAAAAAGTAACATTGAAGATAGAAAACTAGTTAAACAAGCTATAGCAT
TTGATAGACTCTTATCAGAAAAAACGCAAGTTGATCAAAGTAAAATCACA
GCTGAAAGTGAGACAGCTGCGGGGCGATATAACCCTGAAAGTATGGAAAC
GGTTCACAATTACGCCAAGGAATTTGACTTTAAAGAATTGATTGAAAAACT
AGTTGGGCCAACGAATAAGGCAGTCAATGTAGAAGATAAACTTATTTTA
AACAGGTTAATGATGTTATAAATAGTAAACAATTAGCCAATATGAAAGCA
TGGATGATGATTTCTATGCTAGTTGATCAATCAGATTTTCTAGGAGAACAA
AATCGTCAAGCAGCGAGTGCTTTTAAGAATGTTGCGTCTGGTTTGACTCAG
ATTGAATCGAAAGAAAAAATGCTTACACCCAATTAG

MSFMQRKSYLKSM SVLTLTACLISGYVVKDIAM LHAV SASEKKANNVSPREN
LYRAVNDNWLANTKLKQGQTSVNSFSEIEDKLKQLLVSDMAKMASGKIETTN
DEQKKMVA YYKQGMDFKTRDKNGLKPLKPVLQKLEAVSSMKDFQSLAHDF
VMSGFVLPFGLTVETNARDNSQKQLVLRQAPALLES PDQYKKGNKEGEAKLS
AYRTSAMALLKQAGKSNIEDRKL VKQAI AFDRL LSEKTQVDQSKITAESETAA
GRYNPESMETVHNYAKEFD FKELIEKL VGPTNKAVNVEDKTYFKQVNDVINS
KQLANMKAWMMISMLVDQSDFLGEQNRQAASAFKNVASGLTQIESKEKMLT
PN*

Sequence description:

- A] Length: 1095 bp - 365 aa (full length gene)
- B] an GTG (possible ATG start codon located 7 bp further downstream) start codon with an obvious signal peptide. Shine Dalgarno sequence present upstream of the ORF.

ID-90

Clone RS-59

FIG. 1_{CONT'D}

15 / 110

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTTAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTTAAATATTTTTTATAG

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTVAVILSLFK
YFL*

Sequence description:

- A] Length: 174 bp - 58 aa(full length gene)
- B] No obvious signal peptide, but Shine Dalgarno sequence is present upstream of ATG start codon.

ID-91

Clone RS-62 (partial sequence)

ATGCAGGTATTTTTAAATATTGTCAATAAATTCTTTGATCCAGTTATTCATA
TGGGTTTCGGGAGTTGTGATGCTAATTGTCATGACAGGTTTAGCCATGATAT
TTGGAGTGAAGTTTTCTAAAGCACTTGAAGGTGGTAT

MQVFLNIVNKFFDPVIHMSGVVMLIVMTGLAMIFGVKFSKALEGG

Sequence description:

- A] Length:141 bp - 41 aa (partial sequence
- B] Shine Dalgarno sequence present upstream of ATG start codon with a possible signal peptide present

ID-92

FIG. 1_{CONT'D}

16 / 110

Clone RS-69 (partial sequence)

ATGAAAAAGAAAACATTCAGTGCTTATAACTTTTTTAACGGCTCTTATCCTT
TGTCTTTTGACAGTGCTTTTTATCTTTCCATTTTATTGGATTATGACAGGAG
CTTTTAA

MKKKTFSAYNFLTALILCLLTVLFIFPFYWIMTGAF

Sequence description:

- A] Length: 110 bp -36 aa (Partial sequence)
- B] Possible signal peptide with Shine Dalgarno sequence directly upstream of the ATG start codon.

ID-93

Clone RS-70

ATGACTGAGAACTGGTTACATACTAAAGATGGTTCAGATATTTATTATCGT
GTCGTTGGTCAAGGTCAACCGATTGTTTTTTTACATGGCAATAGCTTAAGT
AGTCGCTATTTTGATAAGCAAATAGCATATTTTCTAAGTATTACCAAGTT
ATTGTTATGGATAGTAGAGGGCATGGCAAAAGTCATGCAAAGCTAAATAC
CATTAGTTTCAGGCAAATAGCAGTTGACTTAAAGGATATCTTAGTTTCATT
AGAGATTGATAAAGTTATATTGGTAGGCCATAGCGATGGTGCTAATTTAGC
TTTAGTTTTTCAAACGATGTTTCCAGATATGGTTAGAGGGCTTTTGCTTAAT
TCAGGGAACCTGACTATTCATGGTCAGCGATGGTGGGATATTCTTTTAGTA
AGGATTGCCTATAAATTCCTTCACTATTTAGGGAAACTCTTCCGTATATG
AGGCAAAAAGCTCAAGTTATTTTCGCTTATGTTGGAGGATTTGAAGATTAGT
CCAGCTGATTTACAGCATGTGTCAACTCCTGTAATGGTTTTGGTTGGAAAT
AAGGACATAATTAAGTTAAATCATTCTAAGAAACTTGCTTCTTATTTTCCA
AGGGGGGAGTTTTATTCTTTAGTTGGCTTTGGGCATCACATTATTAAGCAA
GATTCCCATGTTTTTAATATTATTGCAAAAAGTTTATCAACGATACGTTG
AAAGGAGAAATTGTTGAAAAAGCTAATTGA

MTENWLHTKDGSDIYYRVVGQGQPIVFLHGNSLSSRYFDKQIAYFSKYYQVIV
MDSRGHGKSHAKLNTISFRQIAVDLKDILVHLEIDKVILVGHSDGANLALVFQ

FIG. 1_{CONT'D}

FIG. 1 CONT'D

Clone RS-73

TTGAGGGAACTTACTGGAAAATTTCAAGCGATTGCGATAAAAATAAATCTT
GCAGAGTTTTCTAGAGAAAGGAGGTCAGATTTATTGGAGTGGCAAGATCT
AGCGCAGTTACCTGTATCTATTTTTAAAGACTATGTTACAGATGCTCAAGA
CGCGGAAAAACCTTTTATATGGACAGAAGTATTTTAAAGGGAGATTAATCG
CTCAAATCAAGAAATTATTTTGCATATTTGGCCGATGACTAAGACAGTCAT
TCTGGGGATGTTAGATCGAGAATTACCACATTTAGAATTAGCTAAAAAAG
AAATCATCAGTCGTGGTTATGAACCAGTTGTTCCGAATTTTGGAGGTCTCG
CAGTTGTAGCTGATGAAGGAATTTTAAATTTTTCATTGGTTATTCCAGATGT
TTTTGAGAGAAAATTGTCTATCTCAGATGGGTATCTTATAATGGTCGATTTT
ATTAGAAGTATATTTTCGGATTTTATCAACCTATTGAGCACTTTGAAGTA
GAGACCTCCTATTGTCCTGGTAAGTTTGATCTTAGTATAAATGGCAAAAAA
TTTGCTGGCTTGGCTCAGCGCCGTATAAAGAATGGTATTGCGGTATCAATT
TACCTTAGCGTTTGTGGCGATCAAAAAGGGCGGAGTCAAATGATTTTCAGAT
TTTTATAAGATTGGTCTAGGTGATACGGGTAGTCCAATTGCTTATCCAAAT
GTAGATCCTGAAATTATGGCTAATCTATCTGATCTATTAGATTGTCCTATG
ACAGTAGAAGATGTTATTGATCGTATGTTGATTAGCCTTAAACAAGTAGGT
TTTAATGATCGTTTACTGATGATTAGACCCGATTTAGTTGCAGAGTTTGAT
AGATTTTCAGGCTAAGTCTATGGCTAATAAGGGGATGGTGAGCAGAGATGA
ATAA

MRETYWKISSDCDKINLAEFRERRSDLLEWQDLAQLPVSIFKDYVTDAQDAE
KPFIWTEVFLREINRSNQEIILHIWPMKTKTVILGMLDRELPHLELAKKEIISRGYE
PVVRNFGGLAVVADEGILNFSLVIPDVFERKLSISDGYLIMVDFIRSIFSDFYQPI
EHFEVETSYCPGKFDLSINGKKFAGLAQRRIKNGIAVSIYLSVCGDQKGRSQMI
SDFYKIGLGDGTGSPIAYPNVDPEIMANLSDLLDCPMTVEDVIDRMLISLKQVGF
NDRLLMIRPDLVAEFDRFQAKSMANKGMVSRDE*

Sequence description:

- A] Length: 921 bp -307 aa (Full-length gene sequence)
- B] No obvious Shine Dalgarno sequence upstream of the TTG start codon or signal peptide visible. Actual start point may be a further 85 bp downstream (TTG). This start point is preceded by a typical Shine-Dalgarno sequence.

FIG. 1_{CONT'D}

20 / 110

Clone RS-75

ATGACAACCTTACTACGAAGCTATAAACTGGAACGAAATTGAAGATGTTAT
TGATAAATCAACTTGGGAAAACTAACCGAACAATTTTGGCTCGATACAC
GTATCCCTTTATCAAATGACTTAGACGATTGGCGCAAACCTTCCGCTCAAG
AAAAAGATCTTGTTGGCAAGGTTTTTGGAGGCTTAACCCTACTTGATACCA
TGCAATCAGAACTGGTGTGGAAGCTATTCGTGCCGATGTTTCGCACGCCTC
ACGAAGAAGCTGTCTTAAACAATATTCAATTCATGGAATCTGTTACGCTA
AATCTTATTCTTCAATTTTCTCAACTTTAAATACTAAATCAGAAATTGAAG
AAATTTTCGAGTGGACTAATAATAATGAGTTCCTTCAAGAAAAAGCACGT
ATTATCAATGACATTTATGCTAATGGAAATGCCCTTCAAAAAAAGGTGGCT
TCCACCTACCTCGAAACCTTCTTTTTTTATTCTGGCTTTTTTCACACCTCTTTA
CTATTTGGGAAATAATAAGTTAGCAAATGTTGCTGAAATCATTAAATTAAT
TATTCGTGATGAATCTGTACATGGTACTTATATCGGTTACAAATTCAGCTT
GGTTTTAACGAATTACCAGAAGATGAGCAAGAGAATTTTCGTGATTGGAT
GTATGACCTCCTTTATCAGCTGTATGAAAACGAAGAAAAATACACCAAGA
CACTTTATGATGGCGTAGGATGGACTGAAGAAGTTATGACCTTTTTACGCT
ACAATGCTAATAAAGCTCTTATGAATTTAGGACAAGATCCTTTATCCAG
ATACAGCAAATGATGTCAACCCAATTGTTATGAATGGTATTTCAACAGGAA
CATCAAACCATGACTTCTTCTCTCAAGTAGGTAATGGTTACCTACTTGGTA
GCGTTGAAGCTATGCATGATGATGACTATAACTATGGATTATAA

MTTYYEAINWNEIEDVIDKSTWEKLTEQFWLDTRIPLSNDLDDWRKLSAQEK
DLVGKVFGLTLLDTMQSETGVEAIRADV RTPHEEAVLNNIQFMESVHAKSY
SSIFSTLNTKSEIEEIFEWTNNNEFLQEKARIINDIYANGNALQKKVASTYLETF
LFYSGFFTPLYLGNKLANVAEIIKLIIRDES VHGTYIGYKFQLGFNELPEDEQ
ENFRDWMYDLLYQLYENEEKYTKTLYDGVGWTEEVMTFLRYNANKALMNL
GQDPLFPDTANDVNPIVMNGISTGTSNHDFFSQVGNGYLLGSVEAMHDDYN
YGL*

Sequence description:

- A) Length: 960 bp - 320 aa (full length gene)
- B) Shine Dalgarno sequence present upstream of
ATG start codon, but no signal peptide
present.

ID-98

FIG. 1CONT'D

21 / 110

Clone RS-77 (partial sequence)

ATGAATTGGTCACGTATCTGGGAACTCGTAAAAATTAATATCCTTTATTCA
AACCTCAGACTCTATCGGCACTAAGAAAAAGCAAGAAAAGCATCCTAA
AAAAGAATTTTCAGCTTATAAATCCATGTTTAGAAATCAGTTATTTAGAT
TTTGCTCTTTTCAATAATTTATGTATTTCTCTTTGTATCACTTGATTTTAAAG
AATATCCGGGCTATTTACGTTCTACATTGGTATCTTTACACTAGTATCCAT
TATCTACTCTTTTATTGCGATGTACAGTGTTTCTATGAGAGTGACGATGTT
AA

MNWSRIWELVKINILYSNPQTLALRKKQEKHPKKEFSAYKSMFRNQLFQILL
FSIIYVFLFVSLDFKEYPGYFTFYIGIFTLVSIYSFIAMYSVFYESDDV

Sequence description:

- A) Length: 311 bp - 103 aa (Partial sequence)
- B) Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide at
N-terminus.

ID-99

Clone RS-78 (partial sequence)

TAATCTTTTAGTCAACGGAGCAACAGGAAAATTGCAGGCTATGCGACAGA
TATTCACCACATAATTTAGCAGAAGTCATTGATGCTGTCGTGTACATGAT
TGATCACCCTAAAGCTAAATTAGATAAATTAATGGAATTTCTACCTGGTCC
AGATTTTCCAACCTGGCGCTATCATTCAAGGAAAAGATGAAATTCGTAAGG
CATATGAGACTGGTAAGGGGAGAGTAGCGGTTTCGCTCGCGAAGTCTATT
GAAACCTTAAAAGGTGGTAAGAAACAAATTATTGTTACTGAAATTCCTTAT
GAAGTTAAT

SFSQRSNRKLAGYATDIPPHNLAEVIDAVVYMIDHPKAKLDKLMEFLPGPDFPT
GAIQKGKDEIRKAYETGKGRVAVRSRTAETLKGKKQIIVTEIPYEVN

Sequence description:

- A) Length: 312 bp - 104 aa (Partial sequence)
- B) No obvious Shine Dalgarno sequence or a

FIG. 1 CONT'D

FIG. 1 CONT'D

ID-103

Clone 2-11A

ATGGTATTTATGGCAAATAAGAAAAAAACAAAAGGAAAGAAAACCAGAA
GACCTACTAAGGCAGAAATAGAGCGTCAAAGAGCTATTCAAAAGGATGATT
ACTGCTCTTGTTTTAACAATTATTCTCTTCTTTGGTATTATCAGATTAGGTA
TTTTTGGTATTACAGTCTATAACGTCATCCGTTTTATGGTAGGTAGCTTGGC
TACTTATTTATTGCGGCAACTTTAATCTACCTTTATTTCTTTAAATGGTTG
CGAAAGAAAGATAGCTTAGTAGCAGGTTTTTTGATAGCTTCTTTAGGATTA
TTGATTGAGTGGCATGCTTACCTTTTCTCAATGCCTATTTTGAAAGATAAA
GAAATTTTGC GTTCAACTGCTCGATTAAATTGTGTCTGATTAAATGCAATTTA
AAATCACTGTTTTTGCCGGTGGAGGTATGTTGGGTGCTTTGATTACAAGC
CAATTGCTTTTCTCTTTTCTAATATTGGTGCCTATATGATTGGTGTTCTCTTC
ATCATTTTGGGTCTCTTTTTAATGAGTTCTCTGGAAGTTTATGACATCGTCG
AATTTATTAGAGCTTTTAAAAATAAAGTGGCAGAGAAGCACGAGCAAAAT
AAAAAGGAGCGTTTTTGCTAAGCGAGAGATGAAAAAAGCAATCGCTGAACA
AGAGCGCATAGAGCGTCAAAAAGCTGAAGAAGAAGCTTATTTAGCTTCGG
TTAATGTAGACCCTGAAACGGGTGAGATTCTAGAGGATCAAGCTGAGGAC
AATTTGGATGATGCGCTACCACCTGAGGTAAGTGAAACATCAACTCCGGT
ATTTGAGCCAGAGATCCTTGCTTATGAGACATCGCCTCAAAATGATCCTTT
ACCAGTAGAGCCGACAATTTATTTAGAAGACTATGATTCCGCCGATTCTTAA
TATGAGAGAAAATGATGAGGAAATGGTTTATGATTTAGATGATGATGTAG
ATGATAGTGATATAGAAAATGTGCACTTTACACCTAAAACGCACTGGTTT
ATAAATTACCAACGATAGATTTATTTGCACCAGATAAGCCTAAAAATCAAT
CCAAAGAAAAGGATTTAGTCCGAAAGAATATCAGAGTTTTAGAAGAAACA
TTTAGAAGTTTTGGTATCGATGTAAAAGTAGAACGTGCTGAAATTGGACCA
TCAGTTACTAAATATGAAATTAACCAGCAGTTGGAGTTGCTGTAATCGT
ATTTCAAATCTATCTGACGACCTAGCTCTTGCTCTTGCAAGCAAAAGATGTG
CGTATAGAAGCACCAATTCCTGGAAAATCATTAAATAGGTATTGAAGTTCCT
AACTCAGAAATTGCAACGGTTTTCTTCCGCGAACTTTGGGAACAATCTGAT
GCCAATCCTGAAAACCTTTTAGAAGTACCACTAGGAAAAGCTGTAAACGG
CAATGCTCGCAGTTTTAACTTAGCTAGAATGCCGCATCTTTGGTAGCTGG
TTCAACTGGTTCAGGTAAATCTGTGGCAGTTAATGGAATTATTTCAAGTAT
TTTGATGAAGGCACGTCCAGATCAAGTTAAGTTTATGATGATTGATCCCAA
AATGGTTGAATTATCTGTTTATAATGATATTCACATTTATTAATCCCTGTT
GTAACCAATCCGCGTAAAGCAAGTAAGGCACTCCAAAAAGTTGTTGATGA
AATGGAAAATCGATACGAGTTATTTAGCAAAATGGGTGTGCGTAATATAG
CAGGTTATAATACAAAGGTTGAAGAGTTTAATGCTTCCTCTGAGCAAAAAC
AAATGCCTTTGCCTTTAATCGTTGTCATTGTAGATGAATTGGCTGACTTGAT
GATGGTTGCTAGTAAAGAAGTTGAAGATGCTATTATTCGTTTTGGGGCAAAA
AGCACGTGCTGCAGGTATCCATATGATTCTTGCAACTCAACGTCCATCCGT

FIG. 1 CONTD

FIG. 1 CONTD

FIG. 1 CONT'D

FIG. 1 CONT'D

29 / 110

ID-107

Clone 2-54

GAAC TAAATGCAACTCAACCTAATAATAGAACTACCTATATTATACCCGAA
AGCAGTCATTCCATTGCAGAACACAGAGATTCCTGATAGAATCAAAGGG
TTCTTCGGTTGCATTACTTAATAGCGATGAATTTAGAAAAGACAGCGGGAGA
GGATAGAGGTTTTGAAAGGGATAAGTTGAGGTCTTTGGATATCATTCTAA
GGGAGATTTATCGACAAGTAATGTCATAGGTAATACGGACATTGCTAGTC
AGATATCGTTGGGCTTTAAAAAGAATGCGATGCAGGAACACCATCTTACT
AAAACATTCTCTCAAAAGGATGGAAAGTTATCGTCTGTTATAGAGGGGAT
GCTTGCTATTGGCAAAGAGAGAAAGTAGAGAAAGAAATAAAATATAGTGGA
ATTTATGGCAAAAATTAAAGCTAAGGCACACTGCCTTGTTTGCTGTGTG
ATAATTTGAATTTGAAGATATAAAATCTTATTTTCAATATTATTGTCATCT
AAACCATCAGCTCAAATTACCTAAAGGTGCTATACTTTCTGCTAAACAGA
AGTATATAGGGGAGGAGATTTTGGGAGAAAAAATAAAGATAATGTGTTTG
GTTACCGTATCCCCTCATTATTGAAAACCCAAAAAGGAACTTTACTTGCGG
GAGCTGATGAAAGAATTGAGCAAGCTTGTGATTGGGGAAACATAGGAATG
GTTATTCGCCGTAGTGAGGATGATGGTGTCACTTGGGGAAAAAGAGAAAC
TATTGTCAATCTCCGTAATAACCCTAGAGTTCCGCTAGTTACTAGTGGTGA
CTATAGTGGCTCACCTATTAATATGGATATGGCATTAGTTCAAGATACTAG
CTCCAAGACGAAACGTATTTTTTCAATATATGATATGTTTCCAGAAGGAAG
AGGCGTTATTAGTATTGCTAACACACCTGAAAAAGAATATACCCAAATCG
GAGGACAGTCTTATCTTAATTTATATAATAATGGAAAGAAATCGAAGGTTT
TTACTATCCGTGACAAAGGTATTGTATATAATTTTAAAGGGAAAAAGACTG
ATTATCATGTTATAACAGAACTACTAAAAGTGACCATTCAAATCTAGGGG
ATATTTATAAGGGAAAAACAGCTACTTGGAAATATATATTTTACAAAACATA
AAACGTCACCATTTTCGTTTAGCAAAATCAAGCTATGTGTGGATGTCATATA
GCGATGATGATGGTAGGACATGGTCATCACCTAGAGATATAACAGCAAGT
CTTCGTCAGAAAGGCATGAAATTTTTGGGAATAGGACCTGGAAAAGGTAT
AGTTTTAAATGGGGGCCACACGCTGGTCGTATTATTATTCCTGCCTATTCT
ACGAATTGGAAATCTCATCTAAGAGGTTTACAATCTTCACGCCTAATTTAT
TCAGACGACCATGGAAAAACGTGGCATACTGGAAAAGCAGTTAATGATAA
CCGTATACTTTCTAATGGTGAAAAAATTCACCTCCTTAACAATGGATAATAA
AAAAGAACAAAATACAGAATCCGTACCCGTTCAATTGAAAAATGGGGACA
TTAAGTTATTTATGAGGAATCTAACTGGTAACCTAGAAGTAGCCACAAGTA
AAGACGGCGGGGAGACTTGGCAAAACCATGTTAAACGATATAAGGAAATT
CATGATGCTTACGTCCAACCTATCAGCTATTCGCTTTGAGCATGACAAAAAA
GAGTATATTTTATTAGTGAATGCTAATGGGCCAGGGAAGAAGTGCCAAGA
TGGATATGCACGTCTAGCGCAAGTTAATCGAAATGGTAGTTTTAAGTGGTT
ATATCACCATCACATTCAAGATGGTTCGTTTGCTTACAACCTCTGTTCAACA
ACTTAATAATGATCAATTTGGTGTCTTTATGAACATAGAGAAAAACATCA

FIG. 1 CONT'D

30 / 110

AAATAGTTTTACTTTAAATTACAAAGTTTTTAATTGGAGTTTTCTTAGTCAA
AATACAGAGAAGCAAGGCACTTTATGGGAGAAAATGCCAGCAAATTGGCA
TGTTTTGTTTAAATTTTATTTATGA

ELNATQPNRRTTYIIPESHSIAEQQRFLIESKGSSVALLNSDEFKRTAGEDRGF
ERDKLRSLDIIPKGDSTSNVIGNTDIASQISLGFKKNAMQEHHLTKTFSQKDG
KLSSVIEGMLAIGKEKVEKEIKYSGNLWQKLKAKAHCLVCCVDNLNFEDIKS
YFQYYCHLNHQLKLPKGAILSaktevyrGGDFGRKNKDNVFGYRIPSLLKTQ
KGTLLAGADERIEQACDWGNIGMVIRSEDDGVTWVKRETIVNLRNPNRVPL
VTSGDYSGSPINMDMALVQDTSSKTKRIFSIYDMFPEGRGVISANTPEKEYTQI
GGQSYLNLYNNGKKSXVFTIRDKGIVYNFKGKKTDYHVITETTKSDHSNLGDI
YKKGQLLGNIFYTKHTSPFRLAKSSYVWMSYSDDDGRTWSSPRDITASLRQ
KGMKFLGIGPGKGIVLKWGPHAGRIIPAYSTNWKSHLRGSQSSRLIYSDDHG
KTWHTGKAVNDNRILSNGEKIHSITMDNKKEQNTESVPVQLKNGDIKLFMRN
LTGNLEVATSKDGGETWQNHVKRYKEIHDAYVQLSAIRFEHDKKEYILLVNA
NGPGKKCQDGYARLAQVNRNGSFKWLYHHIQDGSFAYNVQQLNNDQFG
VLYEHREKHQNSFTLNYKVFNWSFLSQNTEKQGTLEWKMAANWHVLFKFYL
•

Sequence description:

- A) Length: 2052 bp - 684 aa (partial gene sequence)
- B) N-terminus has yet to be determined

ID-108

Clone 2-61

ATGCCTAAATTAATCGTATCTTTCCTCTGCATTTTATTATCCCTGACTTGTG
TAAACTCTGTGCAAGCTGAAGAACATAAAGATATTATGCAAATTACCCGA
GAAGCCGGATATGATGTTAAAGATATTAATAAACCTAAAGCGTCTATCGTT
ATTGACAATAAAGGTCATATTTTGTGGGAAGATAACGCCGATTTAGAACGT
GATCCCGCTAGCATGTCTAAAATGTTTACTTTATATTTACTATTTGAAGACT
TAGCTAAAGGAAAAACAAACCTCAACACCACAGTGACTGCAACAGAAACA
GACCAAGCCATAAGTAAGATTTATGAAATTAGTAATAACAATATTCATGCT
GGGGTTGCTTATCCTATTCGTGAAGTACTATGACGGCTGTCCCGTCA
TCTAATGTAGCAACTATTATGATTGCTAACCACTTATCACAAAACAATCCT
GACGCCTTTATTAAACGAATCAATGAAACCGCCAAGAACTCGGTATGAC
AAAAACTCACTTTTATAACCCCAAGTGGGGCGGTAGCGAGTGCTTTTAATGG
ACTTTACTCCCCAAAAGAATACGATAACAATGCTACTAACGTTACGACTGC

FIG. 1_{CONT'D}

31 / 110

ACGTGATCTATCAATTTTAACTATCATTTCCTTAAAAAATACCCTGATATA
CTGAACTATACAAAATATCCTGAAGTCAAGGCCATGGTCGGAACCTCTTAT
GAAGAAACATTTACAACCTTATACTACTCTACCCCCGGCGCTAAATTTGGA
TTAGAAGGAGTAGATGGCTTAAAAACTGGTTCTAGCCCTAGCGCTGCTTTT
AATGCCTTAGTTACAGCTAAACGCCAGAATACTCGCTTGATAACTGTGGTT
TTAGGAGTTGGCGATTGGTCAGACCAAGACGGAGAGTACTATCGTCATCC
GTTTGTCAACGCTCTTGTAGAAAAAGGTTTTAAAGACGCTAAAAATATTTT
TTCTAAAACTCCTGTATTAAAAAGCCGTTAAACCTAAAAAAGAAGTTACTAA
AACCAAACTAAATCTATTCAAGAACAGCCTCAAACAAAAGAACAGTGGT
GGACAAAAACAGATCAATTTATCCAATCACATTTTGTATCTATTTTAATTG
TTCTGGGCACCATCGCTAGCCTTTGTCTTTTAGCTGGGATAGTATTACTTAT
AAAGCGCTCTAGATAA

MPKLIVSFLCILLSLTCVNSVQAEHDKDIMQITREAGYDVKDINKPKASIVIDN
KGHILWEDNADLERDPASMSKMFTLYLLFEDLAKGKTNLNTTVTATETDQAI
SKIYEISNNNIHAGVAYPIRELITMTAVPSSNVATIMIANHLSQNNPDAFIKRINE
TAKKLGMTKTHFYNP SGAVASAFNGLYSPKEYDNNATNVTTARDLSILTYHF
LKKYPDILNYTKYPEVKAMVGTPYEETFTTYNYSTPGAKFGLEGVDGLKTGS
SPSAAFNALVTAKRQNTLRITVVLGVGDWSDQDGEYYRHPFVNALVEKGFK
DAKNISSKTPVLKAVKPKKEVTKTKTKSIQEQPQTKEQWWTKTDQFIQSHFVS
ILIVLGTIASLCLLAGIVLLIKRSR*

Sequence description:

- A] Length: 1188 bp - 396 aa (full length gene)
- B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

ID-109

Clone 45

ATGACTGAAAAATATTATAATTGGGCAACGCTTGGAACCGGCGTTATTGCC
AACGAATTAGCCCAAGCACTGGAAGCACGTGGACAAAAATTATATTCTGT
AGCTAATAGAACTTACGACAAAGGACTTGAATTTGCTAACAAATATGGTA
TCCAAAAAGTTTATGATCACATAGATCAAGTATTTGAAGACCCTGAAGTGG
ATATCATTATATCTCTACTCCCCACAATACTCACATCTCATTTTACGAAA

FIG. 1_{CONT'D}

32 / 110

GGCTTTAGCAAATGGTAAGCACGTTCTTTGCGAAAAATCTATTACTTTAAA
TAGTACTGAGCTTAAAGAAGCCATAGATTTAGCCGAAACTAACCATGTTGT
CTTAGCTGAAGCCATGACTATTTTTCATATGCCAATTTACCGCCAATTTAAA
AACATTAGTTGATAGTGGAATAATTAGGACCGTTAAAAATGATTCAAATGA
ATTCGGAAGTTATAAAGAATATGATATGACTAACCGTTTTTTCAGTCGTG
ACCTAGCAGGCGGTGCTTTGCTGGACATTGGTGTTTATGCACTTTCTTGTAT
TCGCTGGTTTATGTCAGAAGCACCTCACAAACATTACCTCTCAAGTTACATT
TGCACCAACAGGGGTTGATGAACAAGTTGGTATCCTACTAACCAACCCAG
CAAATGAGATGGCGACTGTCAGCCTTAGTTTACATGCAAAACAACCTAAA
CGAGCAACTATCGCTTACGATAAAGGCTACATTGAACTTTTTGAATATCCG
CGAGGACAAAAGGCAGTTATTACTTATACTGAGGATGGGCATCAAGATAT
TATCGAAGCTGGCAAACTGAAAATGCTCTCCAATATGAGGTAGCTGATA
TGGAAGAAGCCATTTAGGAAAACTAACACATGTACTTAAACTATACC
AAAGATGTTATGGATATCATGACACAGCTACGTCAAGAATGGGGATTTAC
CTACCCAGAAGAAGAAAAATGA

MTEKYYNWATLGTGVIANELAQALEARQKLYSVANRTYDKGLEFANKYGI
QKVYDHIDQVFEDPEVDIIYISTPHNTHISFLRKALANGKHVLCESITLNSTEL
KEAIDLAETNHVVLAEAMTIFHMPYRQLKTLVDSGKLGPLKMIQMNFGSYK
EYDMTNRFFSRDLAGGALLDIGVYALSCIRWFMSEAPHNITSQVTFAPTGVDE
QVGILLTNPANEMATVSLSLHAKQPKRATIAYDKGYIELFEYPRGQKAVITYT
EDGHQDIIEAGKTENALQYEVADMEEAISGKTNHMYLNYTKDVMDIMTQLR
QEWGFTYPEEEK*

Sequence description:

- A] Length: 984 bp - 328 aa (full length gene)
- B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

ID-110

Clone 2-2

GTGTATTCTCCTGTAAATCTTCTAAAGGAAAAGTGATATTGTTAAAAAGT
GATTTTCTAAAGAGCTTCATAGAAAGGAGAGGAAATATTTGTTTT

MYSPPVKSSKGKVILLKSDFLKSFIERRGNICF

FIG. 1 CONT'D

33 / 110

Sequence description:

- A] Length: 96 bp - 32 aa (partial sequence)
- B] GTG start codon - no obvious Shine-Dalgarno sequence
- Possesses a potential signal peptide

ID-111

Clone 2-3

AAATACTGTATCATTGCAACCTCAAATGCAGGTTTTGGAAACGAAGCATT
ACAGGTGACAGCGATAAAGACTTGAAAATTATGGAACGAATTTCTCCATA
TTCCGTCCAGAATTTCTAAATCGTTTCAATGGTGTTATTGAATTCTCTCAC
CTAAGCAAAGATGACTTAAGCGAAATTGTAGATTTGATGCTTGATGAAGTT
AACCAAACAATTGGCAAAAAAGGAATTGACCTTGTGGTAGATGAAAATGT
TAAATCACACTTAATTGAACTGGGTTATGACGAAGCAATGGGAGTACGTC
CATTGCGCCGTGTCATCGAGCAAGAAATTCGAGATCGCATCACAGACTACT
ATCTCGATCATACAGACGTTAAACACCTAAAAGCTAATTTGCAAGATGGCC
AAATCGTCATTTCTGAAAGATAA

KYCHATSNAFGNEAFTGDSKDLKIMERISPYFRPEFLNRFNGVIEFSLSKD
DLSEIVDLMLDEVNQTIKKGIDLVVDENVKSHLIELGYDEAMGVRPLRRVIE
QEIRDRTDYLDHTDVKHLKANLQDQGQIVISER*

Sequence description:

- A] Length: 429 bp - 143 aa (partial sequence)
- B] N-terminus yet to be elucidated. This gene was not in frame with nuc

ID-112

Clone 2-5

FIG. 1 CONT'D

34 / 110

ATGTCAATGAATTTTTCATTTTACCACAATATTGGTCCTATTTTAATTATG
 GTGTGATGGTAACCATTATGATTTCACATGTGTGTTTTTTTGGAACTAT
 TATAGGCGTGTTAATTGCTTTAGTAAAGCGTACTAATTTACATTTTCTCACA
 ATATTAGCTAATTTCTATGTATGGGTATTTTCGTGGGACACCGATGGTAGTT
 CAAATTATGATTGCTTTCGCATGGATGCATTTTAACAATTTACCAACAATT
 AGCTTTGGTGTTTTAGATTTAGATTTTACACGACTTTTACCTGGTATCATT
 TCATTTCTTAAATAGTGGTGCCTATATTTTCGGAAATTGTACGTGCAGGGA
 TTGAGGCTGTACCATCTGGACAAATAGAAGCAGCTTACTCGTTGGGGATT
 GACCTAAAAATACACTTCGCTATGTTATCTTACCCCAAGCTTTTAAAAATA
 TTTTACCTGCTCTAGGGAATGAATTTATTACAATTATTAAAGATAGTGCTCT
 CCTTCAAACATTGGTGTCTATGGAATTGGAACGGAGCACAATCAGTTGT
 AACGGCTACTTACTCACCAGTTGCACCGTTATTATTTGCAGCATTTTACTAT
 TTAATGTTGACAACGATTCTCTCAGCTTTGTTAAACAAATGGAGAAATAT
 CTTGGGAAAGGGGTAAAAATAGATGGTTGA

MSMNFSFLPQYWSYFNYGVMVTIMISTCVVFFGTIIGVLIALVKRTNLHFLTIL
 ANFYVWVFRGTPMVVQIMIAFAWMHFNNLPTISFGVLDLDFTRLLPGIIISLNS
 GAYISEIVRAGIEAVPSGQIEAAYSLGIRPKNTLRYVILPQAFKNILPALGNEFITI
 IKDSALLQTIGVMELWNGAQSVVTATYSPVAPLLFAAFYYLMLTTILSALLKQ
 MEKYLKGVKIDG*

Sequence description:

- A] Length: 699 bp - 233 aa (full length gene)
- B] Shine-Dalgarno sequence preceded the 'ATG' start codon. Possesses a potential leader peptide sequence.

ID-113

Clone 2-7

ATGAAAGACCTATTACGAAATAGTCTAGAGCAAAGTGGAAATTTAAGTTT
 TCAAGATATGATTTTACATATTCTTGTAGCAGCTTTATTGAGTGTAGTTATT
 TATGTTTCCTATGCTTATACGCATAGTGGAACTGCCTATAGTAAAAAGTTT
 AATGTTTCATTAATGACATTGACGGTCTTGACTGCAACAGTAATGACCGTT
 ATTGGTAATAATGTAGCCTTGTCATTGGGTATGGTTCGGTGCCTTGTCAGTT
 GTTCGTTTTAGGACAGCCATAAAAGATTCAAGAGATACAGTTTATATTTT
 TGGACCATAGTTGTTGGTATCTGTTGTGGTGTTCGGTGAATGTGGTAGCT

FIG. 1 CONT'D

35 / 110

GCATTAGGAAGTAGCGTTATCTTTATCTTATTATGGGTTATGGGACGTGTT
 AAAAACGAGAATCGTATGTTATTGATTGTGAAGTGCATAGAACACTAGA
 AGTTGATTTAGAAGGAATTTTCTTCCAATATTTTGACGGAAAAGCTGTTCA
 GCGTGTTAAAAATTCAACAATACTATTGAAATGATTTTCGAAATCTC
 TAGAAAAGATTACGATAAGCAACTCCATGTAGATAATCAGTTAACTGAAA
 AAGTGTACCAATTGGGAAATATTGATTATTTCAACATTGTTAGCCAAAGCG
 ACGAAATCAATGGGTAG

MKDLLRNSLEQSGNLSFQDMILHILVAALLSVVIYVSYAYTHSGTAYSKKFNV
 SLMTLTVLTATVMTVIGNNVALSLGMVGALSVVRFRITAIKDSRDTVYIFWTIV
 VGICCGVGDYVVAALGSSVIFILLWVMGRVKNNRMLLIVKCDRTLEV DLEGI
 FFQYFDGKAVQRVKNSTTNTIEMIFEISRKYDYDKQLHVDNQLTEKVYQLGNID
 YFNIVSQSDEING*

Sequence description:

- A) Length: 678 bp - 226 aa (full-length gene)
- B) ATG start codon is preceded by a Shine-Dalgarno sequence-Possesses a potential leader peptide sequence

ID-114

Clone 2-8

AAAAATTCATTTTAGATTTCATTTTACGACTATATACTCAGAAGTACCAAAC
 CTAATCCAAGGTTTGAAAAAAGAAAGAAGGAAGTCAGTATGACAAACTAT
 AAAACAACCTTTAAAGATGAGGCTATACGTGTTGAAGAGACAACAAAAGA
 ATCATTTTACGATGTTGATATTGCCTTGTTTTTCAGCTGGTGGATCTATTTCA
 GCAAAGTTTCGCTCCTTATGCAGTAAAGTCTGGAGCAGTTGTAGTAGATAAC
 ACGTCATATTTTCGTCAGAATCCTGATGTTCCACTAGTTGTTCTCCTGAAGTAA
 ATGCTCATGCCATGATTGGTCATAATGGTATCATAGCTTGTCCTCAATTGTTT
 TACTATTCAAATGATGATTGCTTTAGAGCCCATTCGTCAAAAATGGGGGAT
 AGAGCGTGTTATAGTTTCCACCTATCAAGCTGTTTCGGGTTTCAGGTGCACG
 TGCTGTTGAAGAACTAAGGAACAGTTGAGACAAGTTTT

KFILD SFYDYILRSTKPNPRFEKRKKEVSMTNYKNNFKDEAIRVEETTKESFYD
 VDIALFSAGGSISAKFAPYAVKSGAVVVDNTSYFRQNPDPVPLVVPEVNAHAMI
 GHNGIACPN CSTIQMMIALEPIRQKWGIERVIVSTYQAVSGSGARAVEETKEQ
 LRQV

FIG. 1_{CONT'D}

36 / 110

Sequence description:

- A] Length: 499 bp - 165 aa (partial sequence)
- B] N-terminus has yet to be determined

ID-115

Clone 2-9

ATGACAAATGAATTGATAATGCAAGCTTTTGAGTGGTATTTACCTAGTGAT
GGGAATCACTGGAAGAAATTAGAGGAGTCTATATCAGACCTTAAAAAACT
TGGAATTAGTAAAATCTGGTTACCACCAGCATTTAAGGGAAGTAGCAGTG
ATGATGTAGGATATGGTGTTTATGATCTCTTTGATTTAGGAGAATTTGACC
AGAATGGAACAATTAGAACAAAATATGGTAGGAAAGAAGAGTATCTAAA
GCTTATTAAGTCGTAAAGGCAAATGGCATTAAACCGTTTGCAGATATCGT
TCTTAACCATAAAGCCAATGGTGATCATAAAGAAAAATTTCAAGTCATCA
AAGTCAATCCTGAAAATCGTCAAGAAGCATTAAAGTGAACCCTATGAGATT
GAAGGATGGACGGGATTTGATTTCCCAGGTAGACAGGGTGAGTACAATGA
TTTT

MTNELIMQAFEWYLPDGNHWKKLEESISDLKKLGISKIWLPPAFKGTSSDDV
GYGVYDLFDLGEFDQNGTIRTKYGRKEEYLLIKSLKANGIKPFADIVLNHKA
NGDHKEKFQVIKVNPNRQEALSEPYEIEGWTGFDFPGRQGEYNDF

Sequence description:

- A] Length: 456 bp - 152 aa (partial sequence)
- B] ATG start codon is preceded by a Shine-Dalgarno sequence, no leader peptide sequence.

ID-116

Clone 2-10

FIG. 1_{CONT'D}

57 / 110

ATGGAGGTTCTTATGAAGAAAAGTGTTAGTAAGTAGTCTTTTGGTTTTAGGG
ATTACGATAACGTTACAACCAGTAGTTGAGGCTAAGGGGCCAAAAGTAGC
TTATACACAAGAGGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAG
TCACTACTATTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGC
CGATTACTGTTAGTTTTGATATTGATGATACTGCTTTTCAGTAGTCAATA
TTTTCAATATGGTAAAGAATATGTAACCTCGGATCGTTTGATTTTCTTCAT
AAACAAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCAT
TCCCAAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAG
ATAAAATTGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAG
GGCGAGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTTGTA
CCATCTGAT

MEVLMKKVLVSSLLVLGITITLQPVVEAKGPKVAYTQEGMTALSDTNKDKVT
TISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFYQYGKEYVTPGSFDFLHKQKFW
DLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGSMYKEGEVDKTA
KALAKDFKFVPSD

Sequence description:

- A) Length: 516 bp - 172 aa (partial sequence)
- B) ATG start codon is preceded by a Shine-Dalgarno sequence, Possesses a leader peptide sequence.

ID-117

Clone 2-17

ATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGACAAAGATTAGTCGT
CGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTATTATTGTCAGGATGT
GTTTTGGGCCTCAAATTATGATTGAGGGGGTATCAACTCCGAATGTTTCAGC
GCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTAAATTCTTTTCGTAGTTT
AGATCAGCTAACTAGCTTTAAAGAGATTCTTTGGGTTATTGGTCAAAATGT
AGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGTTACTATCCCTAAAG
CCAAGTTTACGGAAATATAAAAGCGTTATATTACTTGCTTTCTTGATGTCTC
TTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTTAATAGATGCTAATCG
GGTTTTTGAAATCGACGATCTATGGACAAATACCTTAGGCGGTCTTTTCGC
CCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCTAACTATTAGAAAATG
A

FIG. 1_{CONT'D}

FIG. 1 CONT'D

39 / 110

B) ATG start codon is preceded by an possible Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-119

Clone 3-7

ATGGAAAAAGAAAAAAATTAGGTCTTTTACCACTAACAATGCTTGTTCATT
GGCTCTCTTATCGGTGGCGGAATCTTTGATTAAATGCAAAATATGAGTTCC
AGAGCCGGTTTGGTACCAATGCTTATTGCTTGGGTAACTACTGCTATCGGG
ATGGGAACTTTCGTTTTAAGTTTTCAAAATTTATCTGAAAAAGGCCGGAC
CTAACAGCTGGAATCTTTAGTTACGCTAAAGAGGGGTTTGGAACTTTATG
GGATTAACTCTGCATGGGGTTATTGGTTATCAGCTTGGCTTGGAAATGTT
GCCTACGCTGCACTCTTATTCAGTTCACCTCGGTTATTTCTTTAAATTCTTTG
GTAATGGAAATAATATCATCTCAATTATTGGAGCAAGTATAGTTATTTGGG
TTGTCCATTTCTTAATTTTAAGAGGTGTTAATACAGCTGCATTTATTAATAC
CGTAGTTACCTTTGCAAAATTAGTACCTGTTATTATTTCTTAATTTACAGCG
TTATTAGCTTTCAAATTTAACATTTTATAGTCTTGATATCTGGGGAAATGGAT
TACATCAATCAATTTTCAACCAAGTCAATTCAACTATGAAAACCGCTGTTT
GGGTATTTATTGGTATTGAGGGCGCCGTTGTCTTCTCAGGTCGTGCTAAAA
AACACTCTGATATTGGTAAAGCAAGTATCCTAGCATTATTCACATATGATTT
CACTTTATGTATTGATTTCTGTTTTATCACTTGGTATCATGTCACGTCCAGA
ACTTGCAAACTTAAAAACACCAGCTATGGCTTACGTTCTAGAAAAAGCTGT
TGGTCACTGGGGTGCTATCTTAGTTAACCTTGGTGTTATCATTTCAGTATTT
GGCGCTATTCTTGCTTGGACTTTATTTGCAGCAGAATTACCATATCAAGCT
GCTAAAGAAGGTGCTTTTCTAAATTTTTTGCAAAAGAAAAATAAAAACAA
AGCTCCAATCAACTCACTCTTAGTCACTAATCTTTGTGTACAAGCATTCTTA
ATCACGTTCTTATTCACACAAAGTGCTTATCGTTTTGGTTTCGCATTAGCAT
CATCTGCTATCTTAATTCCTTATGCTTTTACAGCACTATATCAATTACAATT
CACACTCCGTGAGGATAAGTCAACTCCAGGACATCAAAGAATTTAATTA
TCGGTATCCTCGCTACAATCTATGCTGTTTACCTTATCTACGCTGGTGGTTT
TGATTACTTACTTTTGACAATGATTGCTTATACTCTAGGTATGATTCTCTAT
ATTAAAATGAGAAAAGATGACAAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCAGTGTGAAATTGTTATCC

MEKEKKLGLLPLTMLVIGSLIGGGIFDLMQNMSSRAGLVPMLIAWVITAIGMG
TFVLSFQNLSEKRPDLTAGIFS YAKEGFGNFMGFNSAWGYWLSAWLGNVAY
AALLFSSLGYFFKFFGNGNNIISIIGASIVIWVVFHLILRGVNNTAAFINTVVF
LVPVHIFLISALLAFKFNIFSLDIWGNGLHQSFNQVNSTMKTAVWVFIGIEGAV

FIG. 1_{CONT'D}

40 / 110

VFSGRAKKHSDIGKASILALFTMISLYVLISVLSLGIMSRPELANLKTPAMAYV
LEKAVGHWGAILVNLGVIIISVFGAILAWTLFAAELPYQAAKEGAFPKFFAKEN
KNKAPINSLLVTNLCVQAFLLITFLTQSA YRFGFALASSAILIPYAFTALYQLQF
TLREDKSTPGHQKNLIIGILATIYAVYLIYAGGFDYLLLTMIAYTLGMILYIKMR
KDDKLGVIMVIAVSSVKLLS

Sequence description:

- A) Length: 1356 bp - 452 aa (partial sequence)
- B) ATG start codon is preceded by an possible Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.

ID-120

Clone 3-8

ATGAAATTTGAAAAACGGCAGGTCTATTATGTTGTCATAACATTTGCTATT
TGCTATGCTATACAGGCTTATTGGGGAGCTGTTTCTAATATTTAACTACGC
TTCATAAGGCAATATTTCTTTTTTGGATGGGAGCTGGAATTGCCTATATTAT
TAATATTGTAATGTCAGTCTATGAGCGATTATATATAAAGCTTTTTAAAGG
ATCTAGACTATTAATGGCAATCAAGCGTAGTGTTTCTATGATTTTATCCTAT
GCAACTTTTATTGGTTTAATTGTCTGGCTATTTTCAATTGTCATTCCAGATT
TGATTTCTAGTTTGAGTTCTTTATTGGTTATTGATACCGGAGCACTTGCTAA
ATTGGTTAATAATCTCAATGAAAATAAACAATTTCTGAGGCTTTAAATTA
TATGGGAACAGATAAAGACTTAGTTTCTACTTTAAGTGGTTATAGCCAGCA
GATTTTGAAGCAAGTTTTATCTGTTTTAACAATTTACTAACCTCAGTTTCC
TCTATTGCGGCAACACTTCTGAATGTTTTGTTAGTTTTATTTTTCAATTTA
CGTTTTGGCAAACAAGGAGCAGTTGGGACGTCAATTTAATTGTTAATTGA
TACCTATTTAGGTTCAACAGGCAAAACATTCCATTACGTTTCGTCATATCCTT
CATCAACGTTTCCATGGTTTTTTGTAAGCCAACTTTAGAAGCTATGATTT
TAGGAAGTTTGACGGTTATTGGTATGTTGATCTTCCAATTTCTTATGCTTT
AACAGTTGGGGTTTTAGTTGCTTTTACAGCTCTAATACCGGTTGTGGGAGC
CTACATTGGTGTTACAATCGGTTTCATCTTAATTGCTACTGAATCGTTACT
GAAGCATTCTTGTTGTTCTTTCTTGATCCTTTTACAACAATTTGAGGGAA
ATGTCATTTATCCGAAAGTTGTCGGTGGATCGATTGGACTGCCTTCTATGT
GGGTTTTAATGGCTATTACTATCGGAGGTGCTTTATGGGGGATCTTAGGCA

FIG. 1_{CONT'D}

41 / 110

TGTTACTTGCTGTTTCCTGTTGCAGCTACTATCTATCAGATTGTAAAAGATCA
TATTATCAAGCGACAAACGCTTAGAAATCGTGCACGAACCTATCGTTAA

MKFEKRQVYYVVITFAICYAIQAYWGAVSNILTTLHKAIFPFLMGAGIAYIINI
VMSVYERLYIKLFGKSRLMAIKRSVSMILSYATFIGLIVWLFSIVIPDLISLSS
LLVIDTGALAKLVNNLNENKQISEALNYMGTDKDLVSTLSGYSQQILKQVLSV
LTNLLTSVSSIAATLLNVFVSFISYVLANKEQLGRQFNLLIDTYLGSTGKTFH
YVRHILHQRFHGFFVSQTLEAMILGSLTVIGMLIFQFPYALTVGVLVAFTALIP
VVGAYIGVTIGFILATESLTEAFLFVLILLQQFEGNVIYPKVVGGSIGLPSM
WVLMAITIGGALWGILGMLLAVPVAATYQIVKDHIIKRQTLRNRARTYR*

Sequence description:

- A) Length: 1134 bp - 378 aa (full-length gene)
- B) ATG start codon is preceded by an typical Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.

ID-121

Identical to ID-68, as described in WO 00/06736

ID-122

Clone 3-16

GTGATTACAATTAAAAAGGAATCTGTTATCAAACCTATTGAAGTATGCTTTT
GGCATTATAATGGGATTTATTATCTTAGCTATTGTAATAGGTGGGCTCCTA
TTTGCACTACTACGTTAGTCGTTCTCCGAAATTAACCGATCAAGCTTTAAAA
TCCGTTAACTCTAGTTTGGTTTATGATGGTAATAATAAACTTATTGCCGATT
TAGGCTCAGAAAAGCGTGAAAAGTGTTAGTGCGGATAGCATTCCACTAAAT
TTGGTTAACGCTATCACTTCTATAGAAGATAAACGTTTCTTTAAACATAGA
GGTGTCGATATTTATCGTATTTTAGGTGCAGCTTGGCATAACCTTGTTAGTA
GTAATACGCAAGGTGGTTCAACCCTTGATCAACAGTTGATTAACTGGCTT
ACTTTTCTACCAATAAATCTGACCAAACGTTAAACGTAAATCACAGGAA
GTTTGGCTTGCGCTTCAAATGGAGCGTAAATACACCAAAGAAGAAATTCTT
ACTTTCTATATTAATAAAAGTTTATATGGGAAATGGGAATTATGGTATGAGA

FIG. 1 CONTD

ACAACAGCTAAATCATACTTTGGTAAAGACCTAAAGGAATTATCTATTGCA
CAACTTGCTTTGCTCGCTGGTATTCCTCAAGCACCTACACAATATGACCCTT
ATAAAAACCCAGAATCTGCTCAAACAAGACGTAATACCGTTCTTCAGCAG
ATGTATCAAGATAAAAAACATTTCTAAAAAGGAATACGACCAAGCTGTTGC
AACTCCAGTAACTGATGGCTTAAAAGAATTAAAGCAAAAATCTACTTATCC
AAAATATATGGATAACTACTTAAAACAAGTTATTAGTGAAGTTAAACAAA
AAACTGGTAAAGATATCTTTACTGCTGGGCTAAAAGTGTATACTAATATCA
ACACTGATGCACAAAAACAATATATGACATCTACAACAGTGATACTTAC
ATCGCTTATCCAAACAATGAATTACAAATAGCATCTACCATCATGGATGCG
ACTAATGGTAAAGTCATTGCACAATTAGGCGGGCGTCATCAGAATGAAAA
TATTTTATTTGGGACAAATCAATCTGTCTTAACAGACCGCGATTGGGGTTC
TACAATGAAACCTATCTCAGCTTATGCACCTGCTATTGATAGTGGTGTCTA
TAATTCAACAGGTCAATCATTAAACGACTCAGTTTACTACTGGCCTGGTAC
TTCTACTCAACTATATGACTGGGATCGTCAATATATGGGTTGGATGAGTAT
GCAGACCGCTATTCAACAATCACGTAACGTCCCTGCTGTCAGAGCACTTGA
AGCCGCTGGATTAGACGAAGCAAAATCTTTCCTTGAAAAATTAGGCATAT
ACTATCCAGAAATG

MITIKKESVIKLLKYAFGIIMGFIILAIVIGLLFAYYVSRSPKLTQALKSVNSS
LVYDGNNKLIADLGSEKRESVSADSIPLNLVNAITSIEDKRFFKHRGVDIYRILG
AAWHNLVSSNTQGGSTLDQQLIKLAYFSTNKSDQTLKRKSQEVWLALQMER
KYTKEEILTFYINKVYMGNGNYGMRTTAKSYFGKDLKELSIAQLALLAGIPQA
PTQYDPYKNPESAQTRRNTVLQQMYQDKNISKKEYDQAVATPVTDLKELK
QKSTYPKYMDNYLKQVISEVKQKTGKDIFTAGLKVYTNINTDAQKQLYDIYN
SDTYIAYPNNELQIASTIMDATNGKVIAQLGGRHQENENISFGTNQSVLTD RDW
GSTMKPISAYAPAIIDSGVYNSTGQSLNDSVYYWPGTSTQLYDWD RQYMGWM
SMQTAIQQSRNVPVRALEAAGLDEAKSFLEKLGIIYPEM

Sequence description:

- A) Length: 1386 bp - 462 aa (partial sequence)
- B) GTG start codon is preceded by an
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.

ID-123

Clone 3-17

FIG. 1_{CONT'D}

43 / 110

ATGGCTAATGTATATGATTTAGCAAATGAATTAGAACGTGCTGTTTCGTGCT
TTACCAGAATACCAAGCAGTTTTAACTGCAAAAGCAGCTATTGAAAATGA
TGCGGATGCACAAGTGCTTTGGCAAGACTTTTTGGCTACCCAATCAAAAGT
TCAAGAAATGATGCAATCTGGCCAAATGCCAAGTCAAGAAGAACAAGATG
AAATGTCTAAACTTGGGGAAAAAATTGAATCCAATGACCTTTTAAAAGTTT
ATTTTGACCAACAACAACGGTTGTCTGTCTATATGTCTGATATCGAAAAAA
TTGTCTTTGCACCCATGCAGGACTTGATGTAA

MANVYDLANELERAVRALPEYQAVLTAKAAIENDADAQVLWQDFLATQSK
VQEMMQSGQMPSQEEQDEMSKLGEKIESNDLLKVYFDQQRLSVYMSDIEKI
VFAPMQDLM*

Sequence description:

- A) Length: 336 bp - 112 aa (full length sequence)
- B) ATG start codon is preceded by an
typical Shine-Dalgarno sequence. No obvious
potential leader peptide sequence.

ID-124

Clone 3-26

ATGGCAGAAATCACAGCTAAACTTGTAAAAGAATTGCGTGAAAAATCAGG
TGCAGGCGTTATGGACGCTAAAAAAGCATTAGTAGAAACTGATGGTGACC
TTGATAAAGCGATTGAATTACTTCGCGAAAAAGGTATGGCTAAAGCAGCT
AAAAAAGCAGACCGTGTTGCTGCTGAAGGTTTAACAGGTGTTTATGTTGAT
GGTAACGTTGCAGCAGTTATTGAAGTTAA

MAEITAKLVKELREKSGAGVMDAKKALVETDGDLDKAIELLREKGMAKAAK
KADRVAEGLTGVYVDGNVAIVIEV

Sequence description:

- A) Length: 230 bp - 76 aa (partial sequence)
- B) ATG start codon is preceded by an
typical Shine-Dalgarno sequence. No obvious
potential leader peptide sequence.

FIG. 1_{CONT'D}

45 / 110

ID-127

Clone 3-42

ATGTTAGATATTATCTTATCCGGAATTTGCAAGGATTACTTTGGTCAATTA
TGGCAATTGGCGTGTTTATCACTTTTCGTATCTTAGACATAGCCGATCTCTC
TGCAGAAGGGGCTTCCCTATGGGGGCTGCAGTTTGCGCCTTATGTATCGT
TAA

MLDIILSGISQGLLWSIMAIGVFITFRILDIADLSAEGAFPMGAAVCALCIV

Sequence description:

- A) Length: 158 bp - 52 aa (partial sequence)
- B) ATG start codon is preceded by a possible Shine-Dalgarno sequence. Potential leader peptide sequence.

ID-128

Clone 3-43

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTTAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTAA

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTVAVILSLF

Sequence description:

- A) Length: 161 bp - 53 aa (full-length gene)
- B) ATG start codon is preceded by a possible Shine-Dalgarno sequence. Potential

FIG. 1_{CONT'D}

46 / 110

leader peptide sequence.

ID-129

Clone 3-44

GTGGTAAGTAAATTGAGTTTAACAACGATTTTTGCATTGCTATTTTCATCA
ATGCTAATTTACGCAACACCTCTTATCTTTACAAGTATTGGGGGAACCTTC
TCTGAACGTGGTGGTATCGTCAACGTTGGTTTAGAAGGAATTATGGTAATT
GGAGCTTTCTCAGGCGTTGTATTAA

MVSKLSLTTIFALLFSSMLIYATPLIFTSIGGTFSERGGIVNVGLEGIMVIGAFSG
VVF

Sequence description:

- A] Length: 179 bp - 59 aa (partial sequence)
- B] GTG start codon is preceded by a possible Shine-Dalgarno sequence. Potential leader peptide sequence.

ID-130

Clone 3-46/47

ATGAGAATTATTGCAATAACTGAAAAGGTTATAAAAGAACTGTTTCGTGAT
AAAAGAACACTTGCTATGATGTTTTTAGCACCTATTTAATTATGTTTTGA
TGAATGTTATGTTTTCTGCGAATAGTAATACAAAAGTTAAGATTGGAAC
TTAACGTTAACACGAAGGTCGTTTCAAATTTAGATAATATTAAGCATATTC
AAGTGAGATCATTAAATTTAACTCATCTGCTAAAAAAGCACTCAAATCAA
ATAAAATTGATGCTCTTATTTTCGGAGGACAATAAATCTTATACTGTCTTCT
ATGCGAATACAGATTCTTCAAAGACGACTTTAACAAGACAAGCTTTTAAA
ACCGCTGTTAATACAATGAACAGTAAGGAACTGATTTTCGCAAGTTAAAT
TTAGCTAATAAGAATCCGAACTAGCACAATCCTTACAACTCGCTCCAAA
TATATCAAAGAAAAATATAATTACGGAAATAAAAAATACAGGCTTTTTTGC
AAAAATGATACCAATACTAATGGGATTTATGGTCTTCTTCTTGGTTTTT

FIG. 1 CONT'D

47 / 110

MRIIAITEKVIKELFRDKRTLAMMFLAPILIMFLMNVMFSA NSNTKVKIGTINV
NTKVVS NLDNIKHQVRSFKFNSSAKKALKSNKIDALISEDNKSYTVFYANTDS
SKTTLTRQAFKTA VNTMNSKELISQVKILANKNP KLAQSLQTRSKYIKEKYNY
GNKNTGFFAKMIPILMGFMVFFLVF

Sequence description:

A) Length: 558 bp - 186 aa (partial sequence)
B) ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence. C-terminus has yet to be
determined.

ID-131

Clone 3-48

GTGATTATCGTTATGAGTAAACATCAAGAAATTTTGGAGTACCTAGAAAAT
TTAGCTGTTGGTAAGAGGGTTAGTGTACGCAGTATTTCAAATCATTAA

MIIVMSKHQEILEYLENLAVGKRVS VRSISNHL

Sequence description:

A) Length: 100 bp - 33 aa (partial sequence)
B) GTG start codon is not preceded by a
obvious Shine-Dalgarno sequence. No obvious
leader peptide sequence.

ID-132

Clone 2-c53

FIG. 1 CONT'D

48 / 110

ATGTATAGAGAAATTACCGCTGTCGAACACGATCGCTTTGTGAGCGAATCC
AACCAAACAAACCTACTTCAATCTCTTAATTGGCCCAAAGTAAAAGACAA
CTGGGGTAGTCAATTACTTGGCTTTTTTGGCGGTGAAACCCAAATTGCCAG
CGCTAGTATTCTCATCAAATCACTTCCTCTTGGCTTCTCCATGCTGTATATT
CCGCGTGGACCAATCATGGATTACTCCAATCTAGATATTGTAACCTAAGGTC
CTTAAGGACCTTAAAGCTTTTGGCAAAAAACAAAGAGCTCTCTTTATCAAG
TGTGATCCTCTCATCTATTT

MYREITAVEHDFVSESNTNLLQSLNWPVKVDNWGSQLLGFFDGETQIASA
SILIKSLPLGFSMLYIPRGPIMDYSNLDIVTKVLKDLKAFGKKQRALFIKCDPLI
Y

Sequence description:

- A] Length: 326 bp - 108 aa (partial sequence)
- B] ATG start codon is preceded by an obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-133

Clone 2-c59

ATGGACAAGAAAAAATCTTAGTAACGGGTATTGTGCCTAAAGAAGGTCT
AAGAAAGCTTATGGACCGATTTGATGTTACTTATTCAGAAGATCGCCCAT
TTCACGTGACTATGTGTTAGAGCATTTATCTGAATATGACGGATGGTTACT
CATGGGACAAAAAGGTGATAAAGAGATGATTGATGCAGGTGAAAACCTAC
AAATTATTTCTTT

MDKKKILVTGIVPKEGLRKLMDRFDVTYSEDRPFSRDYVLEHLSEYDGWLLM
GQKGDKEMIDAGENLQIIS

Sequence description:

- A] Length: 215 bp - 71 aa (partial sequence)

FIG. 1 CONT'D

49 / 110

B) ATG start codon is preceded by an obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-134

Clone 2-c62

ATTTGAAAGATGACTACCAAAATATTAGTTTTGGACAGGATCCAGAAGTT
GTTGATTATGCTGGTCTGTTTGAAAAACGCCGTCCAGTTTTAGAAAAAGCA
GTTAAAAATTTCTTGCAAGAAGAGAGCTACGAGAATGCTATCTGATTTT
TTGCAAGAAGAAAAATGGGTAAGTATTTTGGTGAATTTATGGCGATCAA
AGAACATTTTGGTAATAAGGCGCTTCAAGAATGGGATGACAAGGCTATTA
TACGCCGCGAAGAAGAAGCCTTAGCAGGATATCGTCAAAAGCTTAGTGAA
GTGATAAAATATCATGAAGTAACGCAATATTTCTTTTACAAACAATGGTTT
GAGTTAAAAGAATATGCTAATGATAAAGGGATTCAAATTATCCGTGATAT
GCCAATCTACGTTTCTGCCGATAGTGTAGAAGTTTGGACAATGCCTGAACT
GTTT

ISKDDYQNISFGQDPEVVVDYAGLFKRRPVLEKAVKNFLQEERATRMLSDFLQ
EEKWVTDFAEFMAIKEHFGNKALQEWDDKAJIRREEEALAGYRQKLSEVIKY
HEVTQYFFYKQWFELKEYANDKGIQIIGDMPYVSADSVEVWTMPFLF

A) Length: 459 bp - 153 aa (partial sequence)
B) More sequencing is required to determine the
N- and C-termini
enzyme). - Streptococcus pneumoniae (63%)

ID-135

Identical to ID-108 described in WO 00/06736

Clone 2-c63

ID-136

Clone 2-c66

FIG. 1_{CONT'D}

50 / 110

ATGGCAAAACAGAAAAATAACTGGCGCCGTGTTGGAGTTGGTGTCTTAC
 ACTTGCTTCAGTTGCGACTCTTGCTGTCATGTGGAAGTAAATCAGCTTCCCA
 GGATTCTAATGGAGCGATTAAATTGGGCTATTCCAACAGAAATCAATACACT
 AGATTTATCTAAAGTTACAGACACTTACTCAAATCTAGCTATTGGTAACTC
 TAGTAGTAATTTCTTCGCTTAGATAAAGATGGAAAGACAAGACCAGACTT
 GGCTACTAAAGTTGATGTTTCAAAAGATGGCTTAACTTATACAGCTACATT
 ACGTAAAGGCTTGAAGTGGTCAGATGGCAGTAACTTACTGCAAAGGATT
 TTGTTTATTCATGGCAACGTTTAGTTGATCCTAAAACAGCTTCACAATATG
 CTTACCTTGCTGTTGAAGGGCATGTGCTTAATGCCGATAAAATCAACGAAG
 GACAAGAGAAAGACTTGAATAAGCTAGGTGTTAAGGCAGAAGGCGATGA
 CAAAGTTGTTATTACTTTATCTAGTCCGTCTCCGCAATTCATCTACTACCTT
 GCATTCACTAACTTCATGCCACAAAAACAAGAAGTTGTTGAAAAATATGG
 AAAAGATTACGCAACTACTTCAAAAAATACAGTTTACTCAGGACCATATA
 CTGTTGAAGGTTGGAATGGTTCGAATGGTACTTTCACGCTGAAGAAAAAC
 AAAAATTATTGGGACGCTAAAAATGTAAAAACAAAAGAAGTTTCGCATCCA
 GACTGTATAAAAAACCAGATACCGCCGTTCAAATGTATAAACGTGGTGAGT
 TAGATGCAGCTAATATCTCAAATACTTCTGCTATTTATCAAGCTAATAAAAA
 ATAATAAAGATGTACAGATGTTCTAGAAGCGACCACTGCCTATATGGAA
 TATAATACTACTGGTTCTGTGAAAGGGCTTGATAATGTTAAGATTTCGTCGC
 GCCTTAAACTTAGCAACTAACCGTAAAGGAGTTGTTCAAGCAGCCGTTGAT
 ACAGGCTCAAAACCGGCAATTGCTTTTGCACCTACTGGTTTAGCCAAAAACA
 CCAGATGGAAGTATTTGGCAAAATATGTTGCCCCAGGTTATGAATATAAT
 AAAACTGAAGCAGCAAAACTCTTTAGACTA

MAKQKNNWRRVGVGVLTLASVATLAACGSKSASQDSNGAINWAIPTINTLD
 LSKVTDITYSNLAIGNSSSNFLRLDKDGKTRPDLATKVDVSKDGLTYTATLRKG
 LKWSDBGSKLTAKDFVYSWQRLVDPKTASQYAYLAVEGHVNLNADKINEGQEK
 DLNKLGVKAEGDDKVVITLSSPSPQFIYYLAFTNFMPPQKQEVVEKYGKDYAT
 TSKNTVYSGPYTVEGWNGSNGTFTLKKKNKNYWDANKNVKTKEVRIQTVKKPD
 TAVQMYKRGELDAANISNTSAIYQANKNNKDVTDVLEATTAYMEYNTTGSV
 KGLDNVKIRRALNLATNRKGVVQAAVDTGSKPAIAFAPTGLAKTPDGTDLAK
 YVAPGYEYNKTEAAKLFRL

Sequence description:

- A) Length: 1143 bp - 381 aa (partial sequence)
 - B) Shine-Dalgarno sequence precedes ATG codon.
- Possesses a potential leader peptide sequence.

FIG. 1_{CONT'D}

51 / 110

ID-137

Clone 2-c67

TTGAGAGTTTATGAAAATAAAGAAGAGTTGAAAAAAGAAATAAGTAAAAC
ATTTGAGAAATACATTATGGAATTTAATAA
TATTCCAGAGAATCTAAAAGATAAAAGAATTGATGAAGTTGATAGAATC
CAGCAGAAAACCTTTCTTATCAGGTTGGCT
GGACCAACTTGGTTCTTAAATGGGAAGAAGATGAAAGAAAGGGACTTCAA
GTAAAAACACCATCGGATAAATTT

MRVYENKEELKKEISKTFEKYIMEFNNIPENLKDKRIDEVDRTPAENLSYQVG
WTNLVLKWEEDERKGLQVKTPSDKF

Sequence description

- A) Length: 234 bp - 78 aa (partial sequence)
- B) TTG start codon is preceded by a potential Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-138

Clone 2-c70

ATGTCAAAGTTTGATAGTCAGAAAATAAATTACTCCGATTATGAAGTTTGTC
AATATGCGAGGGATTATTGCACTCAAAGATGGCATGCTAGCAATTTTACCA
CTAACAGTTGTTGGGAGTCTCTTTTAAATATTAGGGCAGCTTCCATTT

MSKFDSQKIITPIMKFVNMRGIIALKDGMLAILPLTVVGSFLILGQLPF

Sequence description

- A) Length: 150 bp - 50 aa (partial sequence)
- B) ATG start codon is preceded by a potential Shine-Dalgarno sequence. Possesses a potential

FIG. 1 CONT'D

52 / 110

leader peptide sequence.

ID-139

Clone 2-c71

GAGACCACTTCATCAGTTAAACCAGCAGGAATTGACCGTATCAATCATACC
TCAACACCCCCGAAGAAAACCTACCCCAACATTGCAACGACGCATAGCTT
CAAAGATCGTTGTGATACTTTAGAAAGAATTCACAATGAAGACATTGATGT
TTGTTCTGGATTCAATTTGTGGTATGGGAGAGAGCGATGAGGGGCTCATCAC
ATTAGCTTTCAGACTAAAAGAACTGAACCCCTATTCTATCCCTGTCAATTTT
TACTTGCTGTTGAAGGAACACCTCTTGGAAAATATAACTATTTGACTCCC
ATTAAATGCTTAAAAATTATGGCCATGTTGCGTTTTGTTTTTCCTTTCAAGG
AATTAAGATTAAGTGCTGGACGGGAGGTCCATTTTGAGAATTTTGAATCAT
TAGTCACCTTACTTGTTGACTCAACTTTTTTGGGAAATTACCTAACAGAGG
GGGGTCGCAATCAACATACCGATATTGAATTCTTGGAAAAATTACAATA
AATCATACTAAAAAGGAATTAATTT

ETTSSVKPAGIDRINHTSTPPKKTTPNIAATTHSFKDRCDTLERJHNEDIDVCSGFI
CGMGESDEGLITLAFRLKELNPYSIPVNFLLAEGTPLGKYNLYLTPIKCLKIMA
MLRFVFPFKELRLSAGREVFHFNESLVTLLVDSTFLGNYLTEGGRNQHTDIEF
LEKLQLNHTKKELI

Sequence description:

- A] Length: 535 bp - 178 aa (partial sequence)
- B] N- and C-termini require verification

ID-140

Clone 2-c73

ATGCCGGTTTGGACTGCACAGTCTATTCCAAAGGCATTTTTAGAAAAGCAT
AATACTAAGGAAGGCACCTGGGCAAACTAACCATTCTAAGTGGTTCTTTA
GTATTTTACCAGTTATCTCCTGATGGAGAGGAAATCTCGCGGCATATTTT

FIG. 1_{CONT'D}

53 / 110

GATGCTAGTAGTGATATTCCTTTTGTGATCCACAAGTCTGGCATAAAAGTT
TCGCCGAATAGTCCAGACTTAAGTTGCTATCTAACTTTTTACTGCCAAAAA
GAAGATTACTTCCATAAAAAATATGGTCTCACGCGCACACATTCTGAGGTT
ATCGCCAGTGCACCTCTCTTATCTGAGAAGAGTAATATATTAGACCTTGGG
TGTGGTCAAGGGCGAACTCACTTTATTTATCGCTGCTGGGACATCAAGTG
ACTTCTGTCTGATTCAAACGGACAGAGCCTTGTAGCTTTAGAAAATATGGCA
TTAGAAGAAGAGCTTCCTTACAATATAAAAAAGGTATGATATTAATACTACT
GCTATTGAAGGGCACTATGATTTTATTTTATCAACTGTGGTATTTATGTTTT
T

MPVWTAQSIPKAFLEKHNTKEGTWAKLTILSGSLVFYQLSPDGEEISRHFIDAS
SDIPFVDPQVWHKVSPPNSPDLSCYLTFYCQKEDYFHKKYGLTRTHSEVIASAP
LLSEKSNILDLGCGQGRNSLYLSLLGHQVTSVDSNGQSLVALENMALEEELPY
NIKRYDINTTAIEGHYDFILSTVVFMF

Sequence description:

- A) Length: 563 bp - 187 aa (partial sequence)
- B) N- and C-termini require verification

ID-141

Clone 2c76

ATGACAAAGCAAATAATTGCCATTTGGGCTGAAGATGAAGACCATTTGAT
TGGAGTTAATGGCGGTTTACCATGGAGGCTTCCTAAAGAGTTACATCACTT
CAAAGAAACGACCATGGGGCAGGCTTTGCTTATGGGACGAAAGACCTTTG
ATGGAATGAACCGTCGTGTTTTACCTGGTAGAGAGACAATCATCTTAACAA
AAGATGAACAATTCCAAGCAGATGGAGTGACAGTCCTAAATAGTGTTGAA
CAAGTTATAAAATGGTTTCAGGAACATAATAAGACCTTATTTATTGTAGGT
GGTGCAAGTATTTATAAAGCATTTCTGCCTTATTGTGAAGCAATCATAAAA
ACTAAAGTTCATGGAAAATTCAAAGGTGATACCTATTTTCCTGATGTTAAT
CTATCTGAGTTT

MTKQIIAIWAEDEDHLIGVNGGLPWRLPKELHHFKETTMGQALLMGRKTFDG
MNRRLVLPGRETIILTKDEQFQADGVTVLNSVEQVIKWFQEHNKTLFIVGGASI
YKAFLPYCEAIKTKVHGKFKGDTYFPDVNLSEF

FIG. 1 CONT'D

54 / 110

Sequence description:

- A] Length: 417 bp - 139 aa (partial sequence)
- B] ATG start codon is preceded by a Shine-Dalgarno sequence. No leader peptide sequence

ID-142

Clone 2-c78

TTGTGGCCAAACTGTGCCCCGCTTATTAATAGCACTTTGTTACCAATTGAA
GATATCTTAACATCAGGTGCTCATAGCAACCCTATTTTAATGGGGGTTATA
CTTGGCGGGACAATTGTAGTAGTGGCGACAGCACCACCTTTCTTCTATGGCA
TTGACAGCTATGCTAGGATTAACCGGAATGCCTATGGCTATAGGAGCCTTG
TCTGTCTTTGGTTCGTCATTTATGAATGGTGTACTTTTCCATAAATTAAC
TTGGAAGTCGTAAAGATAATATAGCTTTTGCTGTTGAGCCTCTAACTCAAG
CTGACGTGACTTCAGCTAACCCCTATTCCAATCTATGTCATAATTTTGTGG
TGGTGCAGCTTGTGGTATTTTAATTGCCTTGATGAAATTAGTTAATGATACT
CCTGGAACAGCGACACCAATTGCAGGATTGCTGTCATGTTTGCCTATAAC
CCAATGATAAAAGTACTAATAACCGCTCTAGGTTGTATTATCCTATCTTTA
CTAGCAGGCTATTTTGGAGGCATTGTTTTT

MWPNCAPLINSTLFTIEDILTSGAHSNPILMGVILGGTIVVVATAPLSSMALTA
MLGLTGMPMAIGALSVFGSSFMNGVLFHKLKLGSRKDNIAFAVEPLTQADVT
SANPIPIYVTNFBVGGAAACGILALMKLVNDTPGTATPIAGFAVMFAYNPMIKVL
ITALGCIILSLLAGYFGGIVF

Sequence description:

- A] Length: 540 bp - 180 aa (partial sequence)
- B] N- and C-termini have yet to be elucidated

ID-143

FIG. 1 CONT'D

55 / 110

Clone 2-c80

ATGTTTTTAAGTATAATGGCAGGTGTCATAGCATTGTCTGACAGTTATT
GCCATTCCACGCTTCATTAAGTTTTACCAATTGAAGAAAATTGGCGGGCAA
CAAATGCATGAAGATGTCAAACAACATCTAGCCAAAGCAGGTACGCCGAC
AATGGGAGGAACGGTATTTT

MFLSIMAGVIAFVLTVIAIPRFIKFYQLKKIGGQQMVEDVKQHLAKAGTPTMG
GTVF

Sequence description:

- A) Length: 172 bp - 57 aa (partial sequence)
- B) Shine Dalgarno sequence precedes 'ATG' start codon. Possesses a potential leader peptide sequence.

ID-144

Clone 3-83

ATGAAACCATATTTATCTTTTATTGGTAGAACGTTATTATACTTCGGTATTT
TATTGTTACTAATTACTTTTTTGCATACCTTGGTCGCGGACAAGGCAGTTT
TATTATAA

MKPYLSFIGRTLTYFGILLLLIYFFAYLGRGQGSFIY

Sequence description:

- A) Length: 113 bp - 37 aa (partial sequence)
 - B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.
- This orf is not in frame with nuc

FIG. 1_{CONT'D}

56 / 110

ID-145

Clone 3-86

ATGTCATATTTTAGAAATTACTGGTATCGTTTTGGAGCAATTTTATTTATTA
TTTTAGCAGTAATATTGCTTGTTTTTAGACCTGACTGGTCAATGCTTCACTA
TCTATTGTATTTTACTTTATGGCACTTCTAGCGCATCAATTTGAAGAATAT
CAGTTTCCCGGTGGGGCATCACCTATCATTAACTATGTTGTTTATGATGAA
GAAGAGCTGATGGATTGTTTTCCAGGCAATACTCAGTCTATTATGTTGGTT
AATACTATTGCTTGGTTGCTTTACATTGCTAGTATTGCTTTTCCTCAAGCTT
ATTGGCTTGGATTAGGAGTCATGTTCTTTAGTCTAACGCAGCTCTTGGGTC
ATGGTTTTTCAGATGAATATTAACTTAAAACCTGGTATAATCCTGGTCTAG
CAACGACAGTATTTCTCCTAGTACCAATAGCTTGCGCATACATCTATCAAG
CTAGTGCAGAAGGAATGCTCACTTGGGGAGATTGGCTAGGTGGTTTTATCA
TGTTGATTGTCTGTGTACTAAGTAGCATTATTGCACCTGTACAGCTATTGAA
GGATAAGGAGACCAATTATATTATTAGTCCTTGGCAAATGGACCGTTTTCA
TAAGGTCGTTAATTTTGTAAGGATAAAAAAATAA

MSYFRNYWYRFGAILFIILAVILLVFRPDWSMLHYLLYFYFMALLAHQFEEYQ
FPGGASPIINYVYDEEELMDCFPGNTQSIMLVNTIAWLLYIASIAFPQAYWLG
LGVMFFSLTQLLGHGFQMNILKLTWYNPGLATTVFLVPIACAYIYQASAEG
MLTWGDWLGGFIMLIVCVLTSHAPVQLLKDKETNYIISPWQMDRFHKVVNFV
RIKK*

Sequence description:

- A) Length: 651 bp - 219 aa (full length gene)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.

ID-146

Clone 3-c88

FIG. 1_{CONT'D}

57 / 110

ATGCCACTTACAGCACTTGAAATTAAAGATAAAACATTTTCATCAAAATTT
CGCGGTTATAGCGAAGAAGAAGTT

MPLTALEIKDKTFSSKFRGYSEEEV

Sequence description:

- A) Length: 75 bp - 25 aa (partial sequence)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. No leader peptide

ID-147

Clone 3-90

ATGTCACTTTTTCAAGAAAAAATTGCTTACAATTGCGCTAAAAAGGAAGCG
CTTTATAAAGAGAGTTTAGGACGCTACGCCTTGAGATCAATGCTAGCAGG
GGCTTATTTGACAATGAGTACTGCTGCCGGTATCGTCGCAGCTGATACTAT
TGGTAAAATTTCTCCTGCTCTATCAGGTTTTGTATTTGCTTTCATCTTTAGTT
TTGGACTTATTTATGTTTTAATATTTAATGGTGAATTGGCGACATCTAATAT
GCTTTATCTCACTGCAGGAGCCTATAATAAAAAATATCTCTTGGAAAAAAGC
CATAACAATTTTAATTTATTGTACTTTTTTCAACCTCGTTGGTGCTTGTATA
TTAGCTTGGTTGTTTAA

MSLFQEKIAYNCAKKEALYKESLGRYALRSMLAGAYLTMSTAAGIVAADTIG
KISPALSGFVFVAFIFSGLIYVLIFNGELATSNMLYLTAGAYNKNISWKKAITILI
YCTFFNLVGACILAWLF

Sequence description

- A) Length: 406 bp - 125 aa (partial sequence)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possible leader peptide

FIG. 1_{CONT'D}

59 / 110

ATCCACCAACCAATGGGCGGAACAGGCGGAGGTACACAGCAATCTGATAT
 GGCTATCGCTGCTGAGCATCTTTTAAAAACGCGTCATACTTTAGAAAAAAT
 CTTAGCTGATAATTCTGGTCAATCTATTGAAAAAGTCCATGATGATGCAGA
 GCGTGATCGTTGGATGAGTGCTCAAGAACACTTGATTATGGCTTTATTGAT
 GCTATTATGGAAAAATAATAATTTACAATAATAGATTTAAAAGAGTTGAGTT
 TACCAACTCTTTTTTTATTTGTTGGAATTATGTTATAATCTTAGTAATTACA
 GATATGACGCAGAAAGGAAAAAATTATTGA

MIPVVIEQTSRGERSYDIYSRLLKDRIIMLTGQVEDNMANSIIAQLFLDAQDN
 TKDIYLYVNTPGGSVSAGLAIVDTMNFIKSDVQTIVMGMAASMGTHASSGAK
 GKRFMLPNAEYMIHQPMGGTGGGTQQSDMAIAAEHLLKTRHTLEKILADNSG
 QSIEKVHDDAERDRWMSAQEHLIMALLMLLWKIIYNNRFRKVEFTNSFFICW
 NYVILVITDMTQKGKNY*

Sequence description

- A) Length: 693 bp - 231 aa (full length gene)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. No leader peptide. Significantly, it would appear to have a very hydrophobic C-terminus.

ID-150

Clone 2-c86

ATGAAACCAAAAATTATTGGTGTACTTGGTCTAGGAATATTTGGACAAACA
 CTCGCACAAGAACTAAGTAACCTTTGAACAAGATGTTATTGCTATTGACAGC
 AATCCTGAAAATGTACAAGCTGTCGCCGAAGT
 TGTTACAAAAGCAGCTATCGGAGACATTACTGATTTAGCTTTCCTAAAACA
 CATCGGGATCAGTGACTGTGATACTGTTATTATTGCTACAGGAAACAGTTT
 AGAGAGCTCAGTATTGGCCGTAATGCACTGTAAAAAGTTAGGCGTCCCAC
 AAGTTATTGCTAAAGCTCGAAACCTTGTATACGAAGAAGTACTTTATGAAA
 TTGGTGCTGATTTGGTTATCTCTCCGGAGCGAGAATCTGGGCAAAATGTTG
 CTGCAAACCTCATGAGAAATAAAATTACAGATGTCCTTCAGATTGAATCTG
 ATATTTCTGTCATTGAATTT

MKPKIIGVLGLGIFGQTLAQELSNFEQDVIAIDSNPENVAEAVVTKAAIGDI
 TDLAFLKHIGISDCDVTIIATGNSLE

FIG. 1_{CONT'D}

60 / 110

SSVLAVMHCKKLGVPQVIKARNLVYEEVLYEIGADLVISPERESGQNVAAAN
LMRNKITDVFQIESDISVIEF

Sequence description:

A] Length: 459 bp - 153 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.
This orf is not in frame with nuc

ID-151

Clone 2-c88

GTGCGTTATAGTAAAGAGATTATTCAGTTAGCTATACCAGCTATGATTGAA
AATATCTTACAAATGCTCATGGGAGTAGTTGATAATTATCTAGTGGCTCAG
TTAGGTGTTGTAGCAGTATCAGGTGTTTCAGTTGCTAATAATATAATTACT
ATTTATCAAGCTATTTTTATAGCTTTAGGGGCGAGTATAGCAAGTCTATTG
GCCAAGTCGTTAGCAGGTAGTGAGAAGGATGATGCAATTTTCAGTATGTTCT
CAAGCCATTTTTCTAACATCACTGATAGGGGCGAGTATTAGGAATTATCTCG
ATTGTTTTTGGACAAACTTTCTTT

MRYSKEIIQLAIPAMIENILQMLMGVVDNYLVAQLGVVAVSGVSVANNIITIY
QAIFIALGASIASLLAKSLAGSEKDDAISVCSQAIFLTSLIGAVLGHISIVFGQTFF

Sequence description

A] Length: 330 bp - 110 aa (partial sequence)
B] Putative GTG start codon is preceded by a
typical Shine-Dalgarno sequence. May have a
leader peptide

ID-152

FIG. 1_{CONT'D}

61 / 110

Clone 2-c92

TTGATTAACAAGTATTCGTGCTTTTTGAAGAGGATTCTCCATAATAATACT
CCTTTAATAGTTATCGTGAGAAGTATTTTAAAGAAAAACCGCCAAGGTAG
AGCGACATTTCTGCCTTTAACTACAATAAAACCAAGAGAATTAGCACAAAC
ATTATCTCTCAAAATTACAAAGTTCTCAAGGGTTTTTAGGAATAGCTAGTG
AATTGGTAACCTATGATCAACGCTTGTCAAACATTTTT

MINKYSCFLKRILHNNTPLIVIVRSILKKNRQGRATFLPLTTIKPRELAQHYLSK
LQSSQGFLGIASELVTYDQRLSNIF

Sequence description

A) Length: 240 bp - 80 aa (partial sequence)

B) No obvious Shine Dalgarno sequence precedes the Putative TTG start codon

ID-153

Clone 2-c94

TTGTTGACTCACAAAAATATATTATTAACCATTATATTTGGATTATTTATGA
TTATATTATCAGCATGTGGTATGTCTAATAAGGAAATGGCTGGTATTGATA
ATTGGGAACATTATCAAAAGGAAAAGAAAATTACTATTGGATTGATAAT
ACTTTTGTTCCCTATGGGATTTGAAAGTCGTTCTGGTGACTATACCGGCTTTG
ATATTGATTTAGCTAATGCTGTTTTTAAAGAATACGGTATTTTCAGTGAAAT
GGCAGCCTATTAACCTGGGATATGAAAGAAACTGAACTTAATAATGGTAAT
ATAGACCTTATTTGGAATGGTTATTCAAAAACGGCAGAACGTGCTAAAAA
AGTCGCTTTTACAAACCCATATATGAATAATCATCAAGTAATTGTTACTAA
AACTTCATCACATATTAATAGTATTAAGGATATGAAGGGGAAAAAACTAG
GAGCCCAGTCGGGTTCATCTGGTTTTGATGCTTTTAAACGCTAAACCTGATA
TTTTAAAAAAGTTTGTAAGGAAAAGAAGCAGTTCAATACGATACTTTC
ACTCAGGCTTTGATTGATTTAAAAAATAACCGTATTGATGGTCTTTTGATT
GATGAAGTTTATGCTAACTATTATTAAAGCAAGAAGGAA

FIG. 1_{CONT'D}

62 / 110

MLTHKNILLTIIFGLFMIILSACGMSNEMAGIDNWEHYQKEKKITIGFDNTFV
PMGFESRSGDYTGFDIDLANAVFKEYGISVKWQPINWDMKETELNNGNIDLI
WNGYSKTAERAKKVAFTNPYMNNHQVIVTKTSSHINSIKDMKGKKLGAQSG
SSGFDAFNAKPDILKKFVKGKEAVQYDTFTQALIDLKNNRIDGLLIDEVYANY
YLKQEG

Sequence description

A] Length: 649 bp - 216 aa (partial sequence)
B] TTG start codon is preceded by a possible
typical Shine-Dalgarno sequence. Has a
leader peptide

ID-154

Clone 2-c100

ATGAAAATTTGGAAAAAATAACCTTAATGTTTTCTGCAATTATTTTAACA
ACAGTAATTGCATTGGGAGTCTATGTTGCCTCAGCTTATAATTTTTCGACTA
ATGAATTGTCTAAGACTTTT

MKIWKKITLMFSAIILTTVIALGVYVASAYNFSTNELSKTF

Sequence description

A] Length: 123 bp - 41 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide

ID-155

Clone 2-c1

FIG. 1_{CONT'D}

63 / 110

ATGAAAAAACAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
 ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
 GGAATACCAGGCAGAACAGAAATTTAAGTCATACTTTAAATATATATCAG
 ATAAAAATAACTATTTAGATAATATAAAAAGTTTATTACTTTTCTATAAGTA
 TTTCTAAAGATGTACAAGATAAAGTCAGTGAAACAACAACCTTGTTTCATATA
 GACTAGAAAAGCAAAAGAATCAAGAGTTCATTGGTAATTTTGAACATGAA
 GTTAGTGAATCTAGTCAATATTCAACCGAAGTTAAAAATCAAATACAGTAT
 CCAATCCAGTATAAAGATAATTCAATTCGTTTTACTGAAAAAACACCGTCA
 GAACGTTATGATGAGTTTTGTTTTAGTTCATTTGATTCTTCATTATTAATAA
 AATATAAAATATATGATTACTTACTAAAACATCCCGAAACTGAATTAATAA
 GGTGTTTCCTATAAGATTCTCTATAAATTCTGAAATTGTAGCCCCCTTTATAA
 ATCAATTAAATATAAAAAATCCTAAAAAATCATCTATTTCCGTTACAAAAA
 CGGAAAGTAAAGAATATTATTATACAATCAGTATTGATACTGATTCTGAGA
 TATATTCTATATTCTGAAGGTATTCAT

MKKQRLLLLFGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYFKYISDKN
 NYLDNIKVYYFSISISKDVQDKVSETTTCSYRLEKQKNQEFIGNFEHEVSESSQ
 YSTEVKNQIQYPIQYKDNSIRFTEKTPSERYDEFVFSSFDSSLLKKYKIYDYLLK
 HPETELKGVSYPINSEIVAPFINQLNIKNPKKSSISVTKTESKEYYYTISIDTDS
 EIYSIFEGIH

Sequence description

- A] Length: 687 bp - 229 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. Has a typical leader peptide. C-terminus has yet to be verified

ID-156

Clone 2-c5

ATGACATTTGACACCATTGATCAATTAGCGGTTAATACAGTCCGCACGCTT
 TCTATTGATGCTATCCAAGCAGCAAATTCTGGGCACCCAGGTCTTCCTATG
 GGAGCTGCGCCTATGGCTTATGTGCTTTGGAATAAATTCTTAAATGTAAAC
 CCAAAAACAAGTCGCAATTGGACAAACCGTGACCGTTTTGTACTTTCAGCT

FIG. 1 CONT'D

64 / 110

GGGCATGGTTCAGCTCTTCTTTATAGCCTACTTCATTTAGCTGGCTATGATT
TATCAATTGATGATT

MTFDTIDQLAVNTVRTLSDAIQAANSGLPLMPGAAPMAYVLWNKFLNVNP
KTSRNWTNRDRFVLSAGHGSALLYSLHLAGYDLSIDD

Sequence description

- A] Length: 272 bp - 90 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. No obvious leader peptide

ID-157

Clone 2-c8

ATGAGAACTACTATTTAGAATGATATTTGCTATTCCAAAGTTTATCTTTAGA
TTGATTTGGAATATCATTTGGGGAATATTCAAGACAGTTCTTGTTATTGCG
ATTATTTTATTTGGCTTGTATTACTATGCGAATCACAGTCAATCAGAATTTG
CTAATCAACTTAGTGACATTATTCAGACAGGAAAAACATTTTT

MRTLFRMIFAIPKFIFRLIWNIIWGIFKTVLVIAIHLFGLYYYANHSQSEFANQLS
DIIQTGKTF

Sequence description

- A] Length: 197 bp - 65 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. Possesses a leader peptide

ID-158

FIG. 1_{CONT'D}

65 / 110

Clone 2-c9

ATGTCAAAAAAAAAATAATATTAGGAATTTTATCTCTTTTATCTGTCGTTACTT
 TGGTGGCGTGTGGTTCATCAGACAAACAGCTACAAGATAAAGTTGAGAAA
 AAAGGGAAGTTAGTTTTAGCGGTGAGTCCAGATTATGCTCCCTTTGAGTTT
 MSKKIILGILSLLSVVTLVACGSSDKQLQDKVEKKGKLVLA VSPDYAPFEF

Sequence description

- A] Length: 153 bp - 51 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. Possesses a leader peptide (not in frame with nuc)

ID-159

Clone 2-c10

ATGAAAAATCAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
 ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
 GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTT
 MKNQRLLLLFGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYF

Sequence description

- A] Length: 139 bp - 46 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. Possesses a leader peptide

FIG. 1CONT'D

FIG. 1 CONT'D

67 / 110

MKGLLDLFLVNIARTPAILVALIAIIGLVLQKKGVDPDIVKGGIKTFVGFLVVSEG
AGIVQNSLNPFGKMFHAFHLVGVPNNEAIVAVALTKYGSATALIMLAGMI
FNILIARFTK

Sequence description

- A] Length: 348 bp - 116 aa (partial sequence)
B] ATG start codon is preceded by a potential
Shine-Dalgarno sequence. Possible leader
peptide

ID-162

Clone 2-c21

TTGGTTGGTAAGCCCCAATTACTATTTTTAGATGAACCTACTTCCGGAATG
GATACTTCCACACGTCAACGATTTTGGGAAGCTGGTTGCGACACTAAAAAA
AGAAGGTGACACAATTGTCTATTCTAGTCATTATATCGAAGAGGTAGAAC
ATACAGCTGATAGGATTTTAGTACTTCATAAAGGAAAGTTATTACGCGATA
CAACCCCTTTGCCATGAAGCAAGAAAAAACCGAAAAGTTATTCACCGTT
CCGCTTAGTTATCAAAAATTATTACCTACCTATTTGATTACAGAGTGTGAA
GCCAAGAGTGATAGTATAACGTTTGTTACTGGGGAGGCTGAAACTGTATG
GAAAATACTGGCAGATAATGGTTGTCTATTGAAGCTATTGAGATGACCA
ATAGAACTTTGTTAAATCGTATTTTTGAGACTACTAAGGAGGTAAACATG
AGAATCTTTA

MVGKPQLLFLDEPTSGMDTSTRQRFWKLVA TLKKEGDTIVYSSHYIEEVEHTA
DRILVLHKGKLLRDTPPFAMKQEKTEKLFTVPLSYQKLLPTYLITECEAKSDSI
TFVTGEAETVWKILADNGCPIEAIEMTNRTLLNRIFETTKEVKHENL

Sequence description

- A] Length: 462 bp - 155 aa (partial sequence)
B] Putative TTG start codon is not preceded by
an obvious Shine-Dalgarno sequence. No obvious
leader peptide. N- and C- termini require further

FIG. 1 CONTD

68 / 110

examination.

ID-163

Clone 2-c25

TTGAAAAAATCCAAGAGAAGCCGTAAGGCAGTGACAACAAGTGGTGAGA
AGACTTTACTTGAGGATTTGGCAAAAATGAATTTCTAGACGAAGTCATTA
ATGTTATGGTTTTATATACCTTGAATAAGACAAAATCTGCTAACTTAAATA
AGGCCTATATCATGAAAGTTGCTAATGATTTTGCCTTTCAGAATGTTATGA
CGGCCGAAGATGCTGTGCTTAAAATTCGTGATTTTTCAGATCAAAAAGTAA
GGACTAAACAGAAACGAAGAAGAAACAATCGAATGTTCTGAATGGAGT
AATCCTGATTATAAAGATGAGGTTAGCCCAGAAAAAGAAATTGAATTAGA
ACAGTTT

MKSKRSRKA VTTSGEKTLLDLAKMNFLEVINVMVLYTLNKTKSANLNK
AYIMKVANDFAFQNVMTAEDAVLKIRDFSDQKVRTKTETKKKQSNVPEWSN
PDYKDEV SPEKEIELEQF

Sequence description

A) Length:360 bp - 120 aa (partial sequence)

B) N- and C- termini require verification.

ID-164

Clone 2-c28

ATGACGAATCATATTACTAAACTGATAGAAAATAGCGGAAAAAAATTGAC
AGAAATTAGCGAAGCTACAGATATAGCCTATCCTACACTTTCTGGATACAA
TCAAGGAATCCGCAAACCTAAAAAAGATAATGCTGAAAAATTGGCAAAAT
ACTTTAATGTTTCCGTCGCTTACATTATGGGACTTGATAGCAACCCACATG
CTCCATCAAATCTT

MTNHITKLIENSGKKLTEISEATDIA YPTLSGYNQGIRKPKKDNAEKLAKYFNV
SVAYIMGLDSNPHAPSNL

FIG. 1_{CONT'D}

69 / 110

Sequence description

- A) Length:218 bp - 72 aa (partial sequence)
- B) ATG start codon is preceded by an obvious Shine Dalgarno sequence. No obvious leader peptide.

ID-165

Clone 2-c29

TTGATGAAAAGGAATAAACATTTACCGTTAACAGAACTACCTATTATATT
TTATTAGCTTTGTTTGAGGAAGCGCATGGCTATGCTATTATGAAAAAAGTT
GAAGAAATGAGTGGCGGTGATGTTAGAATAGCCGCAGGGACAATGTACGG
TGCCATTGAAAATTTACTTAAACAAAAATGGATAAAGTCTATCTCAAGTGA
CGATAGAAGAAGAAAAGTTTATATTATTACTGAGACAGGAAAAGAAATAG
TAGAACTTGAAACGAATCGATTAAAGAAAGTTACTTAATACTGCTAATCAGT
TGGGTTTTGGAGGAGATGGTTATGATAAAGTTT

MMKRNKHLPLTETTTYILLALFEEAHGYAIMKKVEEMSGGDVRIAAGTMYG
AIENLLKQKWIKSISSDDRRRKVYIITETGKEIVELETNRLRKLLNTANQLGFG
GDGYDKV

Sequence description

- A) Length:337 bp - 112 aa (partial sequence)
- B) TTG start codon is preceded by an obvious Shine Dalgarno sequence. Actual start codon may ATG that comes immediately after the TTG. Potential leader peptide.

ID-166

FIG. 1_{CONT'D}

70 / 110

Clone 2-c35

CCCATTACTGGTGAGTTAATAGCTGAGAAATTAGGAGTACCAAGAGCAGC
 ACTAAGGTCTGATTTGCGGGTTTTAAGTATGCTAGGTATCATAGATGCAAA
 ACCTAAGGTTGGTTATTTTTATTTAGGACAGTATCATGCTTCAATAGGGAC
 AAGTCATTTTGAAAAGATGACAGTTTCAGAAATTATGGGGATCCTTCTGAC
 AGTTCATCAAAAAGATTCAGTTTATGATGTTATTGTACATATTTTTATGGA
 AGATGCTGGTTGTGCTTTTATCTTGGATGATGATGATTTTCTCTGTGGAGTC
 GTGTCACGTAAAGATTTACTAAAAACCAGTATTGGCGGAGGAGATCTTTCT
 AAAATGCCAATAGGAATGGTGATGACACGTATGCCACACGTGACAACTGT
 TTTAGAAAATGAAAGTCTTTTTGCGGCAGCTGATAAATTAGTGAGCAGAA
 AAGTGGATAGTCTCCCTGTCGTCGTCATGATAAGCAATATCCCGAAAAAT
 TTA

PITGELIAEKLGVPRALRSDLRVLSMLGIIDAKPKVGYFYLGQYHASIGTSHF
 EKMTVSEIMGILLTVHQKDSVYDVIVHIFMEDAGCAFILDDDDFLCGVVSRKD
 LLKTSIGGGDLKMPIGMVMTRMPHVTTVLENESLFAAADKLVSARKVDSLVP
 VRHDKQYPEKF

Sequence description

- A) Length:511 bp - 170 aa (partial sequence)
- B) N- and C-termini to be determined

ID-167

Clone 2-44

TTGGAAGTCATCATGCAATTTATTTATAGTATTATTGGTATTTTATTGGTAT
 TAGGAATTGTGTATGCAATTTCTTTCAATCGTAAGAGTGTTTCTCTAAGTTT
 AATTGGAAAAGCTCTTATCGTTCAATTCATTATTGCGCTAATCTTAGTACGT
 ATCCCACTAGGCCAACAAAGTTGTTAGTGTTGTTTCAACTGGAGTTACTAAA
 GTAATCAACTGTGGTCAAGCTGGTTT

MEVIMQFIYSHIGILLVLGIVYAISFNRSVSLSLIGKALIVQFIILILVRIPLGQQ
 VVSVVSTGVTKVINCGQAG

FIG. 1_{CONT'D}

71 / 110

Sequence description

A) Length:233 bp - 77 aa (partial sequence)
B) TTG start codon is preceded by a possible Shine Dalgarno sequence. Actual start codon may occur further downstream. Potential leader peptide.

ID-168

Clone 2-46

CAACCTAATAAAGCTTTAGAAAAGTGATGAGATTGATATTAATGCTTTCCAG
CATTATAATTACTTAACCAATTGGAATAAAGCAAATAAGACCAATCTTGTT
TCCGTTGCTGAGACATACTTTACTTCCTTTAGATTATACTCTGGTACTAAGA
ACGGTAAAGGTAAATACCAAACAGTTTCTGAAATTCCAAATAAAGCAACT
ATTACTATCCCAAACGATGCAGTTAACGAAAGTCGCTCTCTCTACTTGTTA
CAATCAGCAGGCTTGCTAAAATTGAAAGTATCAGGTGATACATTAGCAAC
AATGTCAGATGTTGTTTCCAATCCTAAATCTTTAGATT

QPNKALESDEIDINAFQHYNLTNWNKANKTNLVSVAETYFTSFRLYSGTKN
GKGKYQTVSEIPNKATITIPNDVNESRSLYLLQSAGLLKLKVSGDTLATMSD
VVSNPKSLD

Sequence description

A) Length:344 bp - 114 aa (partial sequence)
B) N- and C- termini require verification

ID-169

Clone 2-47

ATGAAATGTATAATAAATAATATAAATAAAATAAAAAATGATAATTGAGAT
TTATCATAGAAGGAAAACATTTTTGAAATTAAATAAAATCATATTATCTAC

FIG. 1_{CONT'D}

72 / 110

TGCAGCTCTTACTGCTCTCTTTTTAGGATATAAATAGCGTTACTGCGGATACA
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTAACTGAGGA
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA
ATATACCTTTT

MKCIINNINKIKMIIIEIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYNNNQIELQERIKNLNIPF

Sequence description

- A) Length:264 bp - 88 aa (partial sequence)
- B) There is a Shine-Dalgarno sequence upstream of this sequence. Potential leader peptide sequence

ID-169

Clone 2-47

ATGAAATGTATAATAAATAATATAAATAAAATAAAAAATGATAATTGAGAT
TTATCATAGAAGGAAAACCTATTTTGAAATTAAATAAAATCATATTATCTAC
TGCAGCTCTTACTGCTCTCTTTTTAGGATATAAATAGCGTTACTGCGGATACA
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTAACTGAGGA
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA
ATATACCTTTT

MKCIINNINKIKMIIIEIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYNNNQIELQERIKNLNIPF

Sequence description

- A) Length:264 bp - 88 aa (partial sequence)
- B) There is a Shine-Dalgarno sequence upstream of this sequence. Potential leader peptide sequence

FIG. 1_{CONT'D}

73 / 110

ID-170

Clone RS-58b

TTGGGTGATTATTATGGTAAGAAATATTTTGGTGAGGCAGCTAAAAAAGA
CGTCGAACATATGGCTAAGAAAATCATTAATGTCTATAAAACACGGTTAA
AAAACAACACTTGGTTATC
AGAAAATACAAAAGCAATGGCCATTAAGAACTTGATAACATGAGATTAA
TGATTGGCTATCCAGAAGATTATCCTGATCTTTATCGTCAGTACCAATTTG
ATAGTAAAGCAAGCTTCTTTGAAAACAATGATAACTACAGAAAATTATCG
AACAAGAAAACATTTGAAGAATTTAACCAGTCTAATCAACGTGAACATTG
GCAAATGAGTGCCAATGCTGTAAATGCTTATAATGATCCTAATACCAATTC
CATAGTCTTTCCAGCAGCGATTTTCAATCACCAGTGTACGATAAACTAA
AACAGTTAGTCAAAAATTATGGAGCTATCGGAGCAATTATTGGTCATGAAAT
TTCACACTCATTTGATATTAATGGTATGAAATATGACGAGAAAGGGAATCT
TCACGATTGGTGGACTAAAGAAGATTTAAATCATTATAAGAAATCAACAC
AAGCTATGATTGACCAATGGGATGGCCTTAAAGCAGATGGCGGTAAAGTT
GATGGTAAATTAACCTTTAGCAGAAAATATTGCAGATAATGGTGGTGTTATG
GCATCTCTAGAAGCTCTTAAGACTGAAAAAATCCAACTATAAAGAATTTT
TTGAATCATGGGCAAGTATTTGGCGTCAAAAAGCAACCAAAGAACAAAGT
AAGTCCTCAATTCAGTCAGATGTTTCATGCACCATATGAATTGA >
GAGCTAACATCCCAGTACGTAATTTCCAAGAATTTTATGATGCCTTTGGTG
TTAAAAAAGGCGATTCAATGTATCTAAAACCAGAAAAACGTTTGACACTTT
GTAA

MGDYYGKKYFGGEAAKKDVEHMAKKIINVYKTRLKNNTWLSENTKAMAIAKK
LDNMRLMIGYPDYPDLRQYQFDSKASFFENNDNYRKLSNKKTFEEFNQSNQ
REHWQMSANAVNAYNDPNTNSIVFPAAIFQSPLYDKTKTVSQNYGAIGAIIGH
EISHSFDINGMKYDEKGNLHDWWTKEDLNHYKKSTQAMIDQWDGLKADGG
KVDGKLTLENIADNNGGVMASLEALKTEKIQTIKNFLNHGQVFGVKKQPKNK
VSPQFSQMFMHHMN*

Sequence description:

A) Length: 819 bp - 272 aa (full length gene)
(107 bp of additional DNA sequence (> onwards) is
also included. While not in-frame with the
described orf, it also shares strong homology
with the neutral peptidases.

FIG. 1_{CONT'D}

74 / 110

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-89 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-89 gene sequence. ID-89 and ID-170 together show homology over their combined entire length with the neutral endopeptidases from *Lactococcus* and *Lactobacillus*. Possesses TTG (possible ATG start codon located 13 bp further downstream) start codon with no obvious signal peptide. Shine Dalgarno sequence not immediately obvious. Possibly located further downstream

ID-171

Clone 2-18/22b (Mod2)

ATGACCATGATTACGCCAAGCTTCATTAAGGTATCTCTAGATGAAACAAAT
CGTATGATGCGTATGATATCAGATTTATTAAGTTTATCGCGCATTGATAAT
GAAGTAACGCATTTAGATGTTGAAATGACGAATTTTACAGCTTTCATGACC
TCAATTTTGAATCGATTTGATCAGATTAGAAATCAAAAAACAGTCACAGG
AAAAGTTTATGAAATTGTCAGAGATTATCCTCTTAAGTCAATTTGGGTGGA
AATTGATACAGATAAGATGACTCAAGTGATTGATAACATTTTAAATAATGC
AGTCAAGTATTCACCAGATGGTGGTAAGATTACAGTTAATCTACGCACAAC
TAAACGCAGATGATTTTATCAATATCAGACCAAGGCTTAGGTATTCCCAA
AAAAGATTTACCTCTCATTTTTGATCGTTTTTATCGTGTTGATAAGGCGAGA
AGTCGTCAACAGGGTGGGACTGGACTTGGTTTGTCAATTGCAAAGAAAT
TGTTAAGCAGCATAAGGGATTTATTTGGGCTAAGAGTGAGTATGGTAAAG
GGTCTACTTTTACAATCGTCTTGCCTTATGATAAAGATGCTGTAACCTTATGA
AGAATGGGAGGACGTTGAAGATTAA

MTMITPSFIKVSLEDETNRMMRMISDLLSLSRIDNEVTHLDVEMTNFTAFMTSIL
NRFDQIRNQKTVTGKVYEIVRDYPLKSIWVEIDTDKMTQVIDNILNNAVKYSP
DGGKITVNLRTTKTQMILSISDQGLGIPKKDLPLIFDRFYRVDKARSRQQGGTG
LGLSIAKEIVKQHKGFIAKSEYKGSTFTIVLPYDKDAVTYEEWEDVED*

Sequence description:

- A] Length: 613 bp - 212 aa (full-length gene possibly)
- B] Possible Shine Dalgarno sequence present upstream of a ATG start codon. May not have yet determined the N- portion of this gene. No obvious signal peptide.

FIG. 1_{CONT'D}

75 / 110

ID-172

Clone 2-54balternate (107b)

TTGAAAAAATTATTACTTCTATTCTATTACTTAGTTGCATTTTTTTTATGC
CAACCATCTCTGCTGAATCTTTTAATGCTTCCGCTAAACATGCCTTAGCAGT
TGATTTAGATTCAGGAAAAATCTTGTATGAAAAAGATGCTAACAAACCCG
CTGCTATTGCTTCCTTGACTAAAATAATGACCGTTTATATGGTCTATAAAG
AAATTGATAACGGTAACCTCAAGTGAATACCAAAGTAAATATATCTGAC
TACCCTTATCAACTAACACGCGAATCTGATGCTAGTAATGTTCTTTAGAA
AAAAGGCGCTATACTGTAAACAACCTCGTGGACGCTGCCATGATTTCTAGT
GCTAACAGTGCAGCCATTGCTTTAGCTGAACATATTTTACAGGAAGTAAAAGT
AAATTTGTTGATAAAATGACTGCTCAATTGGAAAAGTGGGGAATTCATGAT
AGCCACCTAGTCAATGCTTCTGGCTTAAATAATAGTATGTTAGGCAATCAC
ATTTATCCAAAATCGTCACAAAACGACGAAAAATAAAATGAGTGCACGTGA
TATTGCTATTGCTGCCTACCATTTGGTCAACGAATATCCTTCCATTCTTAAG
ATTACTAGTAAGTCCGTTGCTAAATTTGATAAAGATATTATGCATTCTTAT
AACTACATGCTACCAGATATGCCTGTCTTTAGACCAGGTATTACAGGTTTG
AAAAGTGGGACAACGGAATTAGCTGGCCAATCTTTTATTGCTACATCTACT
GAAAGTGGAATGAGACTACTCACTGTTATTATGCATGCTGATAAGGCCGAT
AAAGACAAATATGCTCGCTTTACAGCAACTAACTCTCTCTGAACTATATC
ACAAACACCTACGAACCTAACCTTGTATTAGCTAAAGGAGCTGCATATAA
AGGTAAAGAAGCAAGTGTGAGAGACGGAAAAGAACAATCGGTCATCGCT
GTTGCTAAAAACGATTTGAAAGTAGTACAGAAGAAAAATATCACTAAACA
AAATCAGTTAAAAATTAACCTTTAAAAAAGAGCTTACTGCTCCTATTACAAA
AAAAGAGAACCTAGGGAAAGCTTATTACGTTGACCTTAATAAGGTTGGAA
AAGGCTATCTCATAAAGGAACCTAGCGTTCATTTAGTGGCAAAGATAGT
ATTGAGCGCAGTTTCTTCCTCAAAGTGTGGTGAATCATTTTGTGCGCTAC
GTTAACGAAAAACTTTAA

MKKIITSILLSCIFFMPTISAESFNASAKHALAVDLDSGKILYEKDANKPAAIA
SLTKIMTVYVMVYKEIDNGNLKWNTKVNISDYPYQLTRESNASNVPLEKRRT
VKQLVDAAMISSANSAAIALAEHISGTESKFVDKMTAQLEKWGIHDSHLVNA
SGLNNSMLGNHIYPKSSQNDENKMSARDIAIAAYHLVNEYPSILKITSKSVAKF
DKDIMHSYNYMLPDMPVFRPGITGLKTGTTELAGQSFIATSTESGMRLLTVM
HADKADKDKYARFTATNSLLNYITNTYEPNLVLAKGAAYKGKEASVRDGKE
QSVIAVAKNDLKVVQKKNITKQNQLKINFKKELTAPITKKENLGKAYYVDLN
KVGGKGYLIKEPSVHLVAKDSIERSFFLKVWWNHFVRYVNEKL*

FIG. 1 CONT'D

76 / 110

Sequence description:

- A) Length: 1236 bp - 412 aa (full-length gene sequence possibly)
 B) A possible Shine-Dalgarno sequence precedes the putative 'TTG' start codon. (needs further cloning and sequencing to verify N-terminus)

ID-173

Clone 3-60b

ATGACGCTTCGAGAATTAACAATAGAAGAATTTAAAGAACATTCAGGAAA
 TTATGATTCACAATCATTTTTACAAACACCTGAGATGGCTAAACTTTTAGA
 AAAACGCGGCTATGATGTTAGGTATTTGGGATATCAAGTAGAAAATAAAC
 TAGAGATAATCAGTTTATCTTATATTATGCCAGTCACTGGTGGTTTTCAAAT
 GAAAATTGATTCAGGACCAGTTCATTCAAATTCTAAGTATCTAAAACAATT
 TTATAAAGCATTGCAAGGCTATGCCAAATCCAACGGTGTCTAGAATTAAT
 AGTTGAGCCTTTTGATGATTACCAATTATTCAGTTCGGGAGTTCCTAGT
 AATCAGGGAAATGATAATCTGATTGAAGATTTTACCAGTTCAGGTTATCAC
 CATGATGGTTTAACAACCTGGTTTTACTGGTAAATATTTATCTTGGCACTATG
 TTAAAAATTTAGAAGGTGTCACTTCTGAAACGTTACTATCTTCATTCTCTAA
 GACAGGACGAGCTTTGGTTAAGAAAGCAATGTCTTTTGGGAATCAAGGTTT
 GCGTTCCTTAAACGTGATGAGCTACATTTATTTAAAGAGATAACAACCTTCTA
 CGTCAAATAGACGTGATTATATGGATAAGTCCTTAGATTATTATCAAGATT
 TTTACGATAGCTTTGAAGGCAAGGCTGAATTTGTGATTGCCACTTTAAATT
 TTAGAGAATACGACCATAACTTGCAAATAAAAGCTGAAGCATTGGAAAAAT
 AAGCTT

MTLREL TIEEFKEHSGNYDSQSFLQTPEMAKLLEKRGYDVRYLGYQVENKLEI
 ISLSYIMPVTGGFQMKIDSGPVHSNSKYLKQFYKALQGYAKSNGVLELIVEPF
 DDYQLFTSSGVPSNQGNNDNLIEDFTSSGYHHDGLTTGFTGKYLSWHYVKNLE
 GVTSETLLSSFSKTGRALVKKAMSFGIKVRVLKRDELHLFKEITTSTSNRRDY
 MDKSLDYYQDFYDSFEGKAEFVIATLNFREYDHNLQIKAEALENKL

Sequence description

- A) Length: 771 bp - 257 aa (partial gene sequence)
 B) This gene sequence was not identified using the LEEP system. It was identified immediately downstream of the ID-65 gene which was identified by

FIG. 1 CONT'D

77 / 110

LEEP, during cloning and sequence analysis of the full-length ID-65 gene sequence. Sequence Characteristics:
 No obvious leader peptide sequence
 Orf is preceded by a potential Shine-Dalgarno sequence.

ID-174

Clone 2-17b (ID-80b)

TTGTCATTAAGTTTGGTTGCAGTGTTAAATCTTATCCCTCCTAAAATCATGG
 GATCAGTTATTGATGCTATTACAACCTGGAAAATTAACAAGACCACAATTAC
 TATGGAATTTATTAGGTTTGGTTTTGTCAGCTTTAGCTATGTATGGGCTGCG
 TTATATTTGGCGTATGTATATTTAGGGACTTCTTACAAATTAGGCCAAGTT
 GTCAGATACCGTTTATTTGAACATTTACAAAAATGTCTCCTTCTTTTTATC
 AGAAATATCGTACAGGTGATTTAATGGCGCACGCGACCAACGACATCAAT
 TCTCTAACACGTCTTGCAGGAGGAGGAGTTATGTCAGCAGTGGATGCCTCT
 ATCACAGCATTAGTAACGCTTATCACCATGTTCTTTACTATTTTCGTGGCAA
 ATGACATTAATTGCGGTTATCCCTTTGCCCTTAATGGCCTTAGCACTAGTA
 AATTGGGGCGAAAAACCCATGAAACCTTCAAAGAATCTCAGGCAGCCCTT
 TTCAGAATTAATAATAAAGTG

MSLSLVAVLNLIPPKIMGSVIDAITTGKLTRPQLLWNLLGLVLSALAMYGLRYI
 WRMYILGTSYKLGQVVRYRLFHFTHKMSPSFYQKYRTGDLMAHATNDINSLT
 RLAGGGVMSAVDASITALVTLITMFFTISWQMTLIAVIPLPLMALALVNWGEK
 PMKPSKNLRQPFSELNNKV

Sequence description

A) Length: 534 bp - 178 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-80 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-80 gene sequence.

Sequence Characteristics:

No obvious leader peptide sequence

Orf is preceded by a potential Shine-Dalgarno sequence.

FIG. 1_{CONT'D}

79 / 110

AATTTGGACTGTATAACATTTTGGGGATGAATAAACGTCAAGTTGCGCGTG
TAGCTAGTCTAGAGCTGTTTATTATTTATATATTTCTTATTTCTATAGGAAG
TCTGTTTAGTGCTTTTTTTGCTAAATTTATTTATTTAATTTTTGTCAACATTA
TTAACTATCATGCACTAAATCTTAGTTTAAAGTTTATGGCCATTTATTATTTG
TATCGTTATATTTACAGGTATTTTTCTGACTTTAGAAGTTCCAGTTATTCGA
CATGTTCAATTTATCATCCCCATTAAGTCTTTTTAGAAAGAAACAACAGGGA
GAAAAAGAACCAAAAAGGTAATCTTATACCTTGCAATTTTAGCGTTAGTAGCT
ATCGCCATCGCTTATACAATGGCTCTTACTTCAGGTAAAGCACCTGCATTA
GCTGTTATCTATCGTTTCTTCTTTCAGTACTTTTAGTAATTGCTGGTACTT
ATCTTTTTTATATTAGTTTTATGACATGGTACTTAAAAAGGTTGCGTCAAAA
CAAGCATTATTATTATAAATCTGAGCATTTTGTATCAACTTCGCAAATGAT
TTTTCGAATGAAGCAAAATGCAGTAGGGTTAGCAAGTATCACTTTATTAGC
TGTTATGGCTCTAGTTACTATTGCTACAACAGTCTCACTCTATTCAAATACA
CAAAATGTTGTTACCGGACTATTTCCAAAATCAGTAAGTTTATCAATAGAT
AATTCAAAAGGTGACGCGAAAAATATATTTGAAGAAAAGATTTTGAAGAA
ACTAGGTAAGTCATCTAAGGAAGCTATCACTTATAATCAGACAATGATTTC
GATGCCAGTTAGTCAATCAAGTGACTTAATATCACATCTA

MNNMFYLKIAWHNLKHSIDQYIPFLLASLLLYSLTCSTLLILMSAVGRDMGTA
ATVFLGIVVLSIFAVVMEHYSYNILMKQRSSEFGLYNILGMNKRQVARVASL
ELFIIYIFLISIGSLFSAFFAKFIYLIFVNIINYHALNLSLSLWPFHICIVIFTGIFLTLE
VPVIRHVLSSPLSLFRKKQQGEKEPKGNLILAILALVAIAIAYTMALTSBKAP
ALAVIYRFFFAVLLVIAGTYLFYISFMTWYLRQLRQNKHYYYKSEHFVSTSQM
IFRMKQNAVGLASITLLAVMALVTIATTVSLYSNTQNVVTGLFPKSVSLSIDNS
KGDKNIFEKILKKLGKSSKEAITYNQTMISMPVSQSDDLISHL

Sequence description:

- A) Length: 1119 bp - 373 aa (partial gene sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-104 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-104 gene sequence. Possible Shine Dalgarno sequence present upstream of a GTG start codon. Possesses a potential leader peptide sequence

ID-177

FIG. 1_{CONT'D}

80 / 110

Clone 2-5b (ID-112b)

ATGGTTGAGCCAATTATTTCAATACAAGGACTTCATAAAAAGTTTTGGGAAA
AATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTGGT
GGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAAT
GAATCTCTTGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTG
ATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGC
ATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAAT
ATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAG
ACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAA
TGCTTATCCAGCAAGCTTATCTGGAGGACAACAACAACGGATTGCTATTGC
AAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTGATGAACCTACTTCA
GCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTA
GCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCA
CGTGAAGTAGCGGATCGTGTCATTATTATGGATGCAGGGATTATTGTTGAG
CAAGGGACCCCTAAGAAAGTATTTGAGCAGACAAAAGAAATCCGCACAAG
AGACTTCTTAAGTAAAGTATTATAA

MVEPIISIQGLHKSFGKNEVLKGIDLDIHQGEVVVIIGPSGSGKSTFLRTMNLLE
VPTKGTVTFEGIDITDKKNDIFKMREKMGVMVFQQFNLFNMTVLENITLSPIKT
KGLSKLDAQTKAYELLEKVGLKEKANAYPASLSGGQQORIAIARGLAMNPDV
LLFDEPTSAIDPEMVGEVLTVMQDLAKSGMTMVIVTHEMGFAREVADRVIF
MDAGIIVEQGTPKKVFEQTKEIRTRDFLSKVL*

Sequence description:

- A] Length: 735 bp - 244 aa (full length gene)
- B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-112 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence precedes the 'ATG' start codon. No obvious leader peptide

ID-178

Clone 2-5c (ID-112c)

FIG. 1_{CONT'D}

81 / 110

ATGTCTCA^sTATCAAGAGTGGTTAGAAAACGACTCACTCGGTAAAGATATT
 AAGTCAGATTTAGAAGCTATTAAAGGAGATGAATCTGAAATTCAGGATCG
 TTTTACAAAACATTAGAATTTGGAACGGCGGGATTGAGAGGTAAACTTG
 GAGCAGGAACCAATCGTATGAATACTTATATGGTGGGGAAAGCAGCACAA
 GCATTAGCTAATCGATTATTGATCATGGCCCTGAAGCTATTGCACGTGGAA
 TTGCAGTTAGTTATGATGTCCCGTTATCAATCTAAGGAATTTGCAGAATTA
 ACTTGGTCCATTATGGCAGCAAATGGTATTAAAGCCTTATATTTA

MSHMNYKEIYQEWLEND^sLGKDIKSDLEAIKGDSEIQDRFYKTLEFGTAGLR
 KKL^sGAGTNR^sMNTY^sVMVGKAAQALANRLLIMALKLLHVELQLVMMSRYQSKE
 FAELTWSIMAANGIKALYL

Sequence description:

A] Length: 366 bp - 122 aa (partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-112 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence preceded the 'ATG' start codon. No obvious potential leader peptide sequence.

ID-179

Clone 2-5d (ID-112d)

ATGCAACCTGTAAAAGTCGATGAACCTTCTGTTGAAGAAACCATTACTATT
 TTGAAAGGTATCCAAAAAATACGAAGATTATCATCACGTAAAATATAA
 TAATGATGCCATAGAAGCAGCTGCAGTACTATCTAATCGTTATATCCAAGA
 CCGCTTTTACCTGATAAAGCAATAGACTTATTAGATGAAGCTGGTTCTAA
 AATGAACCTAACACTAAATTTTGGTTGATCCAAAAGAAATTGATCAACGTCT
 CATTGAAGCAGAAAATTTAAAAGCGCAAGCGACTCGTGAAGAAGATTACG
 AACGTGCAGCTTACTTCCGTGACCAGATTGCAAAATATAAAGAAATGCAG
 CAACAAAAGGTGCGACGATCAAGATACACCTATTATTACCGAAAAACAAT
 TGAGCACATCATTGAAGAAAAAACGAATATCCCTGTTGGTGATTAAAAAG
 AAAAAGAACAATCTCAATTAATTAATCTCGCAGATGACTTGAAACAGCAT
 GTGATCGGCCAGGATGACGCTGTCATTAAGATTGCAAAAGCTATTCGTCGT
 AATCGAGTTGGTCTTGGTAGCCAAACCGTCCTATTGGTTCTTTTTATTG
 TAGGACCAACCGGTGTTGGTAAAACTGAACTTTCTAAACAAGTACGAATTG
 AGCTCTTTGGTTCAGCTGATAGTATGATTCGTTTTGATATGTCAGAGTACAT
 GGAAAAGCATGCTGTTGCTAAATTAGTCGGAGCGCCTCCAGGATACGTGG
 GATACGAGGAAGCTGGACAATACTGAAAAGGTTTCGTCGAAATCCTTAC
 TCGCTCATCCTTCTAGATGAAATTGAAAAAGCTCATCCCGATGTCATGCAT

FIG. 1_{CONT'D}

82 / 110

ATGTTCTTGCAGGTCCTTGATGACGGTCGATTAAACAGATGGACAAGGAAG
AACTGTTAGTTTTAAAGATACCATTATCATCATGACCTCAAATGCTGGTTC
TGGTAAACTGAAGCAAGTGTGGCTTTGGTGCCTCACGAGAAGGTAGGA
CGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCAT
GCAAGC

MQPVKVDEPSVEETITILKGIQKKYEDYHHVKYNND AIEAAVLSNRYIQDRF
LPDKAIDLLDEAGSKMNLTLNFVDPKEIDQRLIEAENLKAQATREEDYERAA Y
FRDQIAKYKEMQQQKVDDQDTPITEKTIEHIIEEKTNIPVGD LKEKEQSQLNL
ADDLKQHVIGQDDAVIKIAKAIRNRVGLGSPNRPIGSFLFVGPTGVGKTELSK
QLAIELFGSADSMIRFDMSEYMEKHAVAKLVGAPPGYVGYEEAGQLTEKVR R
NPYSLILLDEIEKAHPDVMHMF LQVLDDGRLTDGQGRTVSFKDTIIIMTSNAGS
GKTEASVGF GASREGRTNSSSVPGDPLESTCRHAS

Sequence description:

A] Length: 1070 bp y 356 aa (Partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-112 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence preceded the 'ATG' start codon. No obvious potential leader peptide sequence.

ID-180

Clone 2-7b (ID-113b)

ATGAGAGGGAAGGTTATTTACGGCACAACCCTTATAGGTCTTTTTCTATTC
TTATTTTTCTATTTTTGGATTCCCTAAGCATCACATCGAGAGAATACATCATC
ATCGTATAAAGCAGGTAGATGCGAAGAGTGATTAAACAGGATTTAAAACC
CATTTGCCCATTCAGCATTGATACAAAGCAACAAGTTATTCCTCTTGTT
ACAAAAGAAGGCGGAAAATATGTCAAAGCTAGGGATAATATTAATGTTGA
TATCGAATTACGGGATTCTCCAAGTAGATCACATCATTTATCAGAAAAGCC
GAGAATTAGGACAAAAGGGTTAATATCATATAGAGGAAATTCCTCTCGTT
ACTTTGATAAGAAGTCATTGAAAGTTAAGTTTGTTACTAATAAGTTAAAGG
AAAAGAAGCATCGATTAGCAGGAATGCCTAAAGAATCGGAGTGGGTATTG
CATGGTCCCTTTCTAGACAGAACATTATTAAGAAATTATCTGAGTTATAAT

FIG. 1 CONT'D

83 / 110

ATTGCTGGTGAGATTATGCCTATGCCCCAAACGTTTCGCTACTGTGAGTTAT
TTGTCAATGGTGAGTATCAGGGAG

MRGKVIYGTTLIGLFLFLFFYFWIPKHHIERIHHHRIKQVDAKSDLTGFKTHLP
SIDTKQQVIPLVTKEGGKYVKARDNINVDIELRDSRSHHLEKPRIRTKGLIS
YRGNSSRYFDKKSLLKVKFVTNKLKEKKHRLAGMPKESEWVLHGPFLLDRTLLR
NYLSYNIAGEIMPMPQTFATVSYLSMVSIRES

Sequence description:

- A] Length: 582 bp - 194 aa (Partial gene sequence)
- B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-113 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-113 gene sequence. ATG start codon is preceded by a Shine-Dalgarno sequence-Possesses a potential leader peptide sequence. C-terminus to be determined.

ID-181

Clone 2-17b (ID-117b)

CTTCACATTTTATTGATCACTATCTGACAAATGTTAATCAAACAGCAGTTCT
TATTTTAGTGGGATATTATTCAATGTATGTCTTGCAGACCTTAATTCAATAT
TTTGGGAATCTCTTTTTTGGCGGTGTTTCTTATAGTATTGTTAGAGATATTC
GTAGAGATGCTTTTGCTAATATGGAAAGGCTAGGCATGTCTTATTTTGATA
GGACACCGGCAGGATCTATTGTGTCACGTATTACTAATGATACTGAAGCAA
TATCTGATATGTTTTCGGGTATTTTATCAAGTTTTATCTCGGCGATATTTAT
TTTTACAGTTACTCTGTACACTATGTTGATGCTAGACATTAAACTAACAGG
ACTCGTCGCTCTTTTGTTACCTGTTATCTTTATATTAGTGAATGTCTATCGG
AAAAAATCAGTCACTGTCATTGCTAAAACGAGAAAGTTTACTTAGTGATATC
AACAGTAAATTATCAGAAAGTATTGAAGGAATTC

SHFIDHYLTNVNQTAVLILVGYYSMYVLQTLIQYFGNLFARVSYSIVRDIRRD
AFANMERLGMSYFDRTPAGSIVSRITNDTEAISDMFSGILSSFISAIFIVTLYT
MLMLDIKLTGLVALLLPVIFILVNVRKKSVTVIKTRSLSDINSKLSIESIEGI

FIG. 1 CONT'D

FIG. 1 CONT'D

85 / 110

during cloning and sequence analysis of the full-length ID-120 gene sequence.
ATG start codon is preceded by an typical
Shine-Dalgarno sequence. No obvious leader
peptide sequence

ID-183

Clone 3-11b (ID-121b)

TGGTTAAAAGTAGTGATAGCTTGTATTCCATCTATTTTAATTGCTTTACCAT
TTGATAATTGGTTTGAAGCTCATTTTAATTTTCATGATTCCGATTGCAATAGC
CCTAATCTTTTATGGTTTTGTCTTCATATGGGTTGAAAAACGTAATGCACAC
CTCAAACACAGGTAACCGAATTGGCAAGTATGTCTTACAAGACAGCTTTC
TTGATTGGATGTTTCCAGGTTCTCAGTATTGTTCCGGGAACCAGTCGTTCTG
GAGCTACTATTTTAGGAGCAATTATTATTGGAAGTATGTCGTTCCGGTCGCTG
CTGACTTTACTTTCTTCCTTGCCATCCCAACTATGTTTGGTTATAGTGGACT
TAAGGCGGTTAAATATTTTTTAGATGGTAACGTCTTGAGTTTAGACCAATC
TTTAATACTTTTAGTAGCAAGTCTGACAGCTTTCGTAGTTAGTTTATATGTT
ATTCGTTTCTTGACAGACTATGTCAAACGACACGATTTACCATCTTTGGT
AAGTATCGTATAGTCTTAGGAAGTTTACTCATCCTCTACTGGTTAGTTGTTT
ATTTATTCTAA

WLKVVIACIPSILIALPFDNWFEAHFNFMPIAIALIFYGFVFIWVEKRNAHLKP
QVTELASMSYKTAFLIGCFQVLSIVPGTSRSGATILGAIIGTSRSVAADFTFFLA
IPTMFGYSLKAVKYFLDGNVLSLDQSLILLVASLTAFFVVSLEYVIRFLTDYVKR
HDFTIFGKYRIVLGSLILYWLVVHLF*

Sequence description:

- A) Length: 579 bp - 193 aa (partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-68 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-68 gene sequence described in WO 00/06736. N-terminus has yet to be determined.

ID-184

Clone 3-11c (ID-121c)

FIG. 1_{CONT'D}

86 / 110

ATGGAAATGAAACAAATCAGTGAAACAACACTGAAAATTACAATTAGTAT
GGAAGATTTAGAAGATCGTGGTATGGAGCTGAAAGATTCCTAATCCCTCA
GGAGAAGACTGAGGAATTTTTCTATTCTGTCATGGATGAATTAGACTTGCC
AGAAAACCTTTAAAAATAGTGGTATGTTAAGTTTTTCGAGTAACACCTAAAA
AAGATCGCATTGATGTTTTTGTACAAAGTCTGAATTAAGTAAAGATTAA
ATTTAGAAGAATTAGCAGATTTGGGTGACATTTCAAAAATGTCTCCAGAAG
ACTTTTTTAAAACCTTGGAACAATCGATGTTGGAAAAAGGGGATACGGAT
GCCCATGCCAAATTAGCAGAAATTGAAAATATGATGGATAAAGCAACTCA
AGAAGTAGTTGAGGAAAATGTTTCTGAAGAACAACCTGAAAAGGAAGTAG
AAACGATTGGATATGTTCACTATGTCTTTGATTTTGATAATATTGAAGCTGT
AGTTCGATTTTCAAAACGATTGATTTTCCAATAGAAAGCTT

MEMKQISETTLKITISMEDLEDRGMELKDFLIPQEKTEEFFYSVMDELDLPENF
KNSGMLSFRVTPKKDRIDVFVTKSELSKDLNLEELADLGDISKMSPEDFFKTLE
QSMLEKGD TDAHAKLAEIENMMDKATQEVVEENVSEEQPEKEVETIGYVHY
VFDFDNIEAVVRFSQTIDFPIEA

Sequence description:

- A) Length: 547 bp - 182 aa (Partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-68 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-68 gene sequence. ATG start codon is preceded by an typical Shine-Dalgarno sequence. No obvious potential leader peptide sequence

ID-185

Clone 3-16b (ID-122b)

GGAAACCAACGGCCAGTACAATCGTCAAGGGTAGATTATCCTAAACGTAG
TCGTGCCAAGATTGTAGAAGTTTATTTTAGACAAGCTTCTACTACTGATTA
TTCTGGTGTTTACAAAGGTTACTATATTGACTTTGAAGCCAAAGAAACCCG
GCAGAAAACCTGCTATGCCTATGAAAAATTTTCATGCTCACCAAATAGAGC
ACATGGCAAATGTATTACAGCAAAAAGGGATTGCTTTGTCTTGCTTCATT

FIG. 1 CONT'D

87 / 110

TTTCCACACTTAAGGAAACCTATCTACTCCCTGCTAATGAGTTAATTTTCATT
TTATCAGATTGATAAAGGCAATAAATCAATGCCTATTGATTATATCAGAAA
AAATGGATTTTTCGTAAAGGAGAGTGCCTTTCCTCAAGTCCCTTACTTAGA
TATTATTGAAGAAAAATTATTAGGCGGTGATTACAATTAA

GNQRPVQSSRVDYPKRSRAKIVEVYFRQASTTDYSGVYKGYIIDFEAKETRQ
KTAMPMKNFHAHQIEHMANVLQKGCIFVLLHFSTLKETYLLPANELISFYQI
DKGNKSMPIDYIRKNGFFVKESAFPQVPYLDIIEKLLGGDYN*

Sequence description:

A) Length: 447 bp - 149 aa (partial sequence)

B) This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-122 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-122 gene sequence. N-terminus has yet to be determined

ID-186

Clone 3-17b (ID-123b)

GGATCCTAAAAACGCTAAGGTTTATCAAAAAAATGCTGATCAATTTAGTG
ACAAGGCAATGGCTATTGCAGAGAAGTATAAGCCAAAATTTAAAGCTGCA
AAGTCTAAATACTTTGTGACTTCACATACAGCATTCTCATACTTAGCTAAG
CGATACGGATTGACTCAGTTAGGTATTGCAGGTGTCTCAACCGAGCAAGA
ACCTAGTGCTAAAAAATTAGCCGAAATTCAGGAGTTTGTGAAAACATATA
AGGTTAAGACTATTTTTGTTGAAGAAGGAGTCTCACCTAAATTAGCTCAAG
CAGTAGCTTCAGCTACTCGAGTTAAAATTGCAAGTTTAAGTCCTTTAGAAG
CAGTTCCCAAAAACAATAAAGATTACTTAGAAAATTTGGAAACTAATCTTA
AGGTACTTGTCAAATCGTTAAATCAATAG

DPKNAKVYQKNADQFSDKAMAIAEKYKPKFKAASKYFVTSHTAFSYLAKR
YGLTQLGIAGVSTEQEPSAKKLAEIQEFVKTYKVKTIFVEEGVSPKLAQAVAS
ATRVKIASLSPLEAVPKNNKDYLENLETNLKVLVKS LNQ*

Sequence description:

FIG. 1 CONT'D

88 / 110

A] Length: 433 bp - 144 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-123 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-123 gene sequence. N-terminus has yet to be determined

ID-187

Clone 3-46/47 (ID-130b)

ATGAAAAAAGTCATCGATTAAAAAACTACAAAAAGCATACGCCTCAGA
AACTGTTTTAAATAATATTAATTTGGAGGTGTTAAAGGAGAAATAATTGG
ATTAATAGGACCCTCTGGAGCAGGGAATCTACCTTGATTAATACTATGCT
TGGCATGGAAAAAGCAGATAAGGGAACAGCTCTTGTTCTTGATACTCAA
TGCCAGATCGTAATATTTTAAATCAAATTGGCTATATGGCTCAATCTGATG
CCTTACACGAGTCTTTAACTGGCTTAGAAAAATTTATTATTCTTTGGAAAA
TGAAAGGTATTCAAAAACTGAATTAACAGCAGATAACTCATATTTCT
AAAGTAGTAGATCTAGAAAACCAACTTGATAAATTTGTCTCAGGTTACTCA
GAAGGTATGAAAAGACGGCTTTCTCTAGCCATCGCCCTACTTGGAAACCCC
ACAGTTTTAATCCTAGATGAACCTACCGTTGGAATTGATCCATCCTTGAGG
AGAAAAATCTGGCAAGAGCTAATTAATATTAAGGATGAAGGACGTTCTAT
CTTTATTACAACCCACGTTATGGATGAAGCAGAATTAACAAGTAAGGTTGC
ACTACTATTACGTGGAAACATTATTGCCTTTGATACTCCATTACATTTAAA
AAAACAATTTAATGTGAGTACTATTGAGGAAGTTTTCTTAAAGCTGAAGG
AGAATAA

MKKVIDLKKLQKAYASETVLNNINLEVFKGEIIGLIGPSGAGKSTLIKTMLGME
KADKGTALVLDTPMPDRNINLQIGYMAQSDALHESLTGLENLLFFGKMKGIQ
KTELKQQITHISKVVDLENQLDKFVSGYSEGMKRRLSLAIALLGNPVLILDEP
TVGIDPSLRRKIWQELINIKDEGRSIFITTHVMDEAELTSKVALLLRGNIIAFDTP
LHLKKQFNVSTIEEVFLKAEGE*

Sequence description:

A] Length: 717 bp - 239 aa (Possible full-length sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-130 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-130 gene sequence. ATG start codon is preceded by a possible

FIG. 1_{CONT'D}

89 / 110

Shine-Dalgarno. No obvious potential leader
peptide sequence

ID-188

Clone 3-83b (ID-144b)

ATGGTACAAATGATACATGATATGATTAAACAATTGAGCATTGCTGAG
ACACAAGCTGATTTTCCAGTGTATGATATTTAGGGGAAGTCCATACTTAT
GGACAACCTAAAGTAGACTCTGACTCTCTAGCTGCTCATATTGATAGCCTA
GGCCTTGTTGAAAAATCACCTGTCTTAGTATTCCGGTGGTCAAGAATATGAA
ATGTTGGCGACATTTGTTGCTTTAACAAAGTCAGGGCATGCTTATATACCG
GTTGACCAACACTCTGCTTTGGATAGAAATACAGGCTATTATGACAGTTGCT
CAACCAAGCCTTATCATTTC AATTGGTGAATTTCTCTTGAAGTTGATAAT
GTCCCAATCCTAGACGTTTCTCAAGTTTCAGCTATTTTGAAGAAAAGACT
CCTTATGAGGTAACACATTCTGTTAAAGGTGATGATAATTACTATATTATT
TTCACCTTCAGGGACTACTGGTTTACCAAAAAGGTGTGCAAATTTACATGAC
AATTTATTGAGCTTTACAAATTGGATGATTTCTGATGATGAGTTTTCAGTTC
CTGAAAGACCGCAAATGTTGGCTCAACCC

MVQMIHDMIKTIEHFAETQADFPVYDILGEVHTYGQLKVDSDSLAAHIDSLGL
VEKSPVLVFGGQEYEMLATFVALTKSGHAYIPVDQHSALDRIQAIMTVAQPSL
IISIGEFPLEVDNVPILDVSQVSAIFEKTPYEVTHSVKGDDNYIIFTSGTTGLP
KGVQISHDNLLSFTNWMISDDEFSVPERPQMLAQP

Sequence description:

- A) Length: 592 bp - 197 aa (partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-144 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-144 gene sequence. Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. No obvious leader peptide sequence
This orf is not in frame with nuc

ID-189

FIG. 1_{CONT'D}

90 / 110

Clone 3-86b (ID-145b)

ATGGAAAATCATCGTTATGAAGATGAAGGTAAATTCCAGCGTAAGATGAC
CAGTCGTCATCTCTTTATGTTATCGCTAGGTGGTGTATCGGGACTGGGCTT
TTCTTGAGTTCAGGTTATACCATTGCACAGGCTGGTCCGCTTGGAGCTGTG
CTGTCTTATTTGATTGGTGCCGTTGTGGTTTATTTGGTCATGCTATCACTTG
GGGAATTGGCGGTTGCCATGCCGGTGACGGGGTCATTCCACACTTATGCCA
CTAAGTTTATCAGTCCTGGAACAGGTTTTACTGTTGCTTGGCTATATTGGAT
TTGTTGGACGGTCGCCTTGGGGACTGAATTTTATAGGTGCTGCCATGCTGAT
GCAGCGCTGGTTCCCAAATGTGCCGGCTTGGGCATTTGCTTCCTTTTTTGCC
CTTGATTTTTGGTTTAAATGCTCTTAGCGTACGCTTTTTTGCAGAAAGCAG
AGTCTTCTCTCAAGTATTAAGGTTATTGCTATCATTATCTTTATTATCTTG
GGCTTAGGTGCTATGTTTGGTCTAGTTTCCTTTGAAGGTCAGCACAAAGGCT
ATTCTCTTCACTCATCTGACTGCCAATGGTGCCTTCCCAAATGGTATCGTTG
CAGTTGTCTCAGTCATGTTGGCTGTTAACTATGCCTTCTCTGGTACTGAGTT
AATTGGTATTGCGGCTGGTGAAACGGATAATCCCAAAGAAGCTGTACCAA
GGGCTATTAAAACGACAATCGGTGCGTTGGTTGTTTTCTTTGTACTGACAA
TTGTTGTCCTAGCTTCGCTATTGCCAATGAAAGAGGCAGGCGTATCCACAG
CACCATTGCTTATGTCTTTGACAAGATGGGAATCCCTTTTACGGCGGATA
TCATGAACCTTCGTTATCTTGACAGCCATCCTGTCTGCTGGTAACTCAGGTCT
CTACGCATCAAGCCGATGCTCTGGTCCCTTGCCAATGAAGGTATGTTGTC
AAAATCTGTTGTGAAAATCAATAAACACGGGTGTCCTCAATGCGTGCTCTTCT
CTTGTCATGGCAGGAGCAGTGCTGTGCTCTTTTCAAGTATTTACGCTGC
AGACACAGTTTATCTAGCCTTGGTTTCAATCGCGGGCTTTGCTGTTGTTGTC
GTATGGCTAGCCATTCCAGTCGCACAAATCAATTTCCGCAAGGAATTC

MENHRYEDEGKFQRKMTSRHLFMLS LGGVIGTGLFLSSGYTIAQAGPLGAVL
SYLIGAVVVYLVMLSLGELAVAMPVTGSFHTYATKFISPGTGFTVAWLYWIC
WTVALGTEFLGAAMLQMQRWFPNVPWAFASFFALVIFGLNALSVRFFAEAES
FFSSIKVIAHIFILGLGAMFGLVSFEGQHKAILFTHLTANGAFPNGIVAVVSVM
LAVNYAFSGTELIGIAAGETDNPKEAVPRAIKTTIGRLVVFFVLTIVVLASLLPM
KEAGVSTAPFVDVFDKMGIPFTADIMNFVILTAILSAGNSGLYASSRMLWSLA
NEGMLS KSVVKINKHGVPMRALLSMAGAVLSL FSSIYAADTVYLALVSIAGF
AVVVVWLAIPVAQINFRKEF

Sequence description:

A) Length: 1126 bp - 393 aa (partial gene)

FIG. 1 CONT'D

91 / 110

sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-145 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-145 gene sequence. Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a possible leader peptide sequence.

ID-190

Clone 3-94b

TCAGAAAATGCAGAGGCAGCAACGGTTGCCACAACTTGGTTACCAAAGG
AGCTAATGTCATTATCGGACCAGCAACATCGGGTGCAGCTGCATCTTCAAC
TCCAAAAGTAAATGCAGCAGCAGTTCCAATGATTGCACCTGCTGCGACAC
AAGACAATTTAGTCTATGGTTCTGATGGAAAAACCTTAAATCAGTATTTCT
TCCGAGCTACTTTTGTGCGATAATTATCAAGGAAAGCTATTGTCTCAGTATG
CTACAGACAACCTTAAAGCTAAAAAAGTTGTTCTATTTTATGATAATTCAT
CAGATTACTCAAAGGGGGTAGCAAAATCATTTAAGGAAAGTTATAGTGGA
AAAATTGTTGATAGTATGACATTCTCCGCTGGTGATACTGATTTCCAAGCG
TCATTGACTAAGTTGAAAGGGGAAAGAATATGATGCTATTGTGATGCCAGG
TTACTATACCGAGACAGGATTAATAGTTAAGCAAGCGCGTGATTTAGGTAT
CTCTAAACCGGTTCTTGGGCCTGATGGTTTTGATAGTCCGAAATTTGTGCA
ATCGGCAACACCTGTAGGAGCTTCAAACGTTTATTATTTGACAGGTTTCAC
TACACAAGGATCAACCAAAGCTAAAGCT

SENAEAATVATNLVTKGANVIIGPATSGAAASSTPKVNAAAVPMIAPAATQD
NLVYGSDGKTLNQYFFRATFVDNYQGKLLSQYATDNLKAKKVVLFDNSSD
YSKGVAKSFKESYSGKIVDSMTFSAGDADFQASLTKLKGKEYDAIVMPGYT
ETGLIVKQARDLGISKPVLPDGFDFSPKFVQSATPVGASNVYYLTGFTTQGST
KAKA

Sequence description

A] Length: 637 bp - 231 aa (partial sequence)

FIG. 1_{CONT'D}

92 / 110

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-149 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-149 gene sequence. N- and C-termini have yet to be determined

ID-191

Clone 2-c94b (ID-153b)

TTGGGACTTAAAGACCATGCTTTAGTCTATCCATTTTCATTATCTGGGGGG
CAAAAGCAACGTGTCGCACTAGCTCGTGCGATGATGATTGATCCACAGATT
ATTGGTTATGATGAGCCAACTAGCGCTCTTGATCCAGAGTTGCGTCAAGAA
GTAGAAAACTAATTTTACAAAATAGAGAAACAGGTATGACACAAATTGT
AGTAACACATGATCTTCAATTTGCTGAAAGTATATCTGATACGATTCTCAA
AATTAATCCTAAGTAG

MGLKDHALVYPFSLSGGQKQRVALARAMMIDPQIIGYDEPTSALDPELRQEV
EKLILQNRETGMTQIVVTHDLQFAESISDTILKINPK*

Sequence description

A] Length: 270 bp - 90 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-153 gene which was identified by LEEP, during cloning and sequence analysis of the ID-153 gene sequence.

N-terminus has yet to be determined

ID-192

Clone 2-c1b (ID-155b)

ATGACTAATATCTCAGATGTTCCAAAAGCTATTAGAACACAGGCACAGTAT
GTTCTCTTGGGAATGAGAGTTATGGATCAGTCGGTATTACCGAAAACATAT
AATTCAAAAGAACCTTATTTGAAACCAGATATGATTTATATTCATGATAGA

FIG. 1_{CONT'D}

FIG. 1 CONT'D

94 / 110

GGTTTACCTTGATGGTGAAGGAGCACGTGGCACGATGTTGTCTATTGCTTT
TGCAAACAAAGGACAACACCAAGATACGGGTGCAAAGATGATTCATAATG
CCCCCATACTAGTTCATCCATTGTCTCTAAATCAATTGCTAAGGGTGGGG
GAAAAGTTGATTATCGAGGTCAAGTGACATTTAATAAAGATTCCAAAAA
TCAGTGTACATATAGAATGTGACACCATATTGATGGATGATATTTCAAAA
TCAGATACCATAACCGTTTAATGAGATTCATAATTCACAGGTTGCTTTAGAG
CATGAAGCAAAGGTGTCTAAGATTTCTGAAGAGCAACTGTACTACTTGATG
AGTCGAGGTTTATCTGAAGCTGAAGCAACAGAAATGATTGTTATGGGGTTT
GTTGAGCCCTTTACGAAAGAATTACCAATGGAATATGCGGTAGAGTTAAA
TCGTTTAATTTCTATGAAATGGAAGGTTGAGTTGGTTAA

MHACRSTLEDLGEYKFGFHDDVKPIYSTGKGLNEAVIRELSAAKGEPEWMLD
FRLKSLETFNKMPMQTWGADLSIDFDDIYYQKASDKPARDWDDVPEKIKE
TFERIGIPEAERAYLAGASAQYSEVVYHNMKEEYDKLGIVFTDTSALKEYP
ELFKKYFAKLVPPTDNKLAALNSAVWSGGTFIYVPKGVKVDIPLQTYFRJNE
NTGQFERTLIIVDEGASVHYVEGCTAPTYSSNSLHAAIVEIFALDGAYMRYTTI
QNWSDNVYNLVTKRATAKKDATVEWIDGNLGAKTMMKYPSVYLDGEGARG
TMLSIAFANKGQHQDTGAKMIHNAPHTSSSIVSKSIKGGGKVDYRGQVTFN
KDSKKSVSHECDTILMDDISKSDTIPFNEIHNSQVALEHEAKVSKISEEQLYYL
MSRGLSEAEATEMIVMGFVEPFTKELPMEYAVELNRLISYEMEGSVG*

Sequence description:

- A] Length: 1411 bp - 469 aa (Possible full-length gene)
- B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-72 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-72 gene sequence. No obvious Shine Dalgarno sequence upstream of TTG start codon (insufficient sequence data). N terminus needs verification.

ID-194

Clone 3-1b (ID-81b)

ATGATAGAATTCTTTTCTAATATCAGAACAGAGATTCCGCAGATGCCTTTA
CTTATCCATAGTTTGATTTTATCTGTCTTACCTTTTCTGATGTGGCTGACTTT
GGTTAATAGAGATAAGCCTTTGTATAAACTATTTGGAGTATCCTTTTAGG
ACTTCAGTTAATTACGATTTATACTTGGTTTTTCTGGGCAAATTGCCTTTA

FIG. 1 CONT'D

95 / 110

TCTGAAAGTCTTCCCCTTTACCATTTGTCGAATAGGCATGTTTGTCCGGTCTCT
TA

MIEFFSNIRTEIPQMPLLIHSLILSVLPFLMWLTLVNRDKPLYKTIWSILLGLQLI
TIYTWFFWAKLPLSESLPLYHCRIGMFVGLL

Sequence description

A) Length: 261 bp - 87 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-81 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-81 gene sequence. Sequence Characteristics: Possesses a potential leader peptide sequence. Orf is preceded by a potential Shine-Dalgarno sequence.

ID-195

Clone RS-55b

AAGCTTGTGCAAAGTATTAAGAGATAGGATTAGCTAATGCGCATTATTATTA
GCTGTTGCTCCGACAGGGTCAATCAGTTATCTTTCTTCTTGTACTCCGAGCC
TTCAACCGGTTGTATCACCTGTCGAAGTACGCAAGGAAGGAGCACTGGGG
AGGGTTTATGTAGCTGCTTATAAGATTGATGCAGATAATTATGTCTACTAC
AAAAAAGGAGCTTATGAAGTGGGATCTGAGGCGATTATCAATATTGCAGC
TGCCGCTCAAAAACACATTGATCAAGCTATTTTCGTTAACGCTTTTCATGAC
AGATCAAGCAACTACGCGAGATTTAAATAAAGCCTATATTCAAGCATTTA
AACAAAAATGTGCCTCTATTTATTATGTACGAGTGAGACAGGACATCCTAG
AAGGTAGCGAGAGTTATGATGATATGCTGGATGATTTCACTTCATCGGACT
TAGAAGACTGTCAATCCTGCATGATTAA

>KLVQSIKEIGLANAHLAVAPTGSISYLSSCTPSLQPVVSPVEVRKEGALGRV
YVAAYKIDADNYVYYKKGAYEVGSEAIINIAAAAQKHIDQAISLTLFMTDQAT
TRDLNKAYIQAFKQKCASIYYVRVRQDILEGSESYDDMLDDFTSSDLEDCQSC
MI*

Sequence description:

FIG. 1_{CONT'D}

96 / 110

A] Length 486 bp - 162 aa (Partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-87 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-87 gene sequence. N-terminus to be determined.

ID-196

Clone RS-59(ID-90b)

GTGAGGACATATATTACAACTTGAATGGACATTCAATCACTAGTACAGC
ACAAATAGCTCAAAACATGGTAACAGATATAGCAGTAAGCTTAGGTTTC
GTGAGCTGGGAATACATTCTTATCCGATTGATACTGATTCTCCTGAGGAAA
TGAGTAAGCGTTTAGATGGAATCTGTTCCGGACTTAGAAAAAATGATATTG
TCATATTTTCAGACACCTACATGGAACACTACAACCTTTTGATGAAAAATTAT
TTCACAAATTAAAAATATTTGGTGTAAGATTGTTATTTTATACATGATGT
TGTACCGCTAATGTTTGATGGAAATTTTATTTGATGGATAGAACTATAGC
TTATTATAATGAAGCAGATGTTTAATAGCCCCTAGTCAAGCAATGGTCGAT
AAGCTT

MRTYITNLNGHSITSTAQIAQNMVTDIAVSLGFRELGIHSPIDTDSPEEMSKRL
DGICSLRKNDIVIFQTPTWNTTTFDEKLFHKLKIFGVKIVIFIHDVVPLMFDGN
FYLMDRTIAYYNEADVLIAPSQAMVDKL

Sequence description:

A] Length: 414 bp - 138 aa(partial gene)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-90 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-90 gene sequence. No obvious signal peptide, but a possible Shine Dalgarno sequence is present upstream of ATG start codon. C-terminus has yet to be determined.

ID-197

Clone RS-59c (ID-90c)

FIG. 1^{CONT'D}

FIG. 1 CONT'D

98 / 110

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence.
N-terminus has yet to be determined

ID-199

Clone RS-70c (ID-93c)

ATGAAATTAAGTGTCCTTGATTATGGGCTTATTGATTATGGAAAACTGCA
AGTGATGCAATACAAGAAACGATTCTTTTATCACAAGAGGCGGAGCAACT
AGGCTATCATCAATTTTGGGTGGCTGAACATCACGGTGTTAAGGCATTTCAG
TATTAGCAATCCAGAATTAATGATAATGCATTTGGCTAACCAGACTAAATC
TATCAAAATTGGCTCTGGAGGTATAATGCCTCTGCACTATAGTAGTTTTAA
ACTCGCGGAGACTCTCAAGACATTAGAGACATGTCATCCGAATCGAGTAA
GTATTGGTTTAGGAAATTCAGTAGGGACAGTTAAAGTTTCAAATGCACTTC
GTAGCTTACATAAAGCACATGATTACGAAGAGGTACTGGAGGAATTGAAG
TCATGGCTTATTGATGAATCATCCAGTAAGGAACCATTAGTTCAACCGACT
CTTTCTAGCTTCCCAGACTTATATGTGTTGGGGAGTGGTCAAAAATCAGCT
TATTAGCGGCTAAACTTGGCTTAGGCTTACCTTCGGTGTTTTTCTTTTA
TGGACAAAGACCCATTGACAGAAGCTAAA

MKLSVLDYGLIDYGKTASDAIQETILLSQEAELGYHQFWVAEHHGVKAFFSIS
NPELMIMHLANQTKSIKIGSGGIMPLHYSSFKLAETLKTLETCHPNRVSIGLGN
SLGTVKVSNALRSLHKAHDYEEVLEELKSWLIDESSSKEPLVQPTLSSFPDLYV
LGSGQKSAYLAACLGLGFTFGVFPFMDKDPLTEAK

Sequence description:

A] Length: 588 bp - 196 aa (partial)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence. No obvious signal peptide, but Shine Dalgarno sequence upstream of the ATG start codon.

FIG. 1_{CONT'D}

99 / 110

nucS1

Bgl II Eco RV

5'-cgagatctgatatctcacaaacagataacggcgtaaataag -3'

nucS2

Bgl II Sma I

5'-gaagatcttccccgggatcacaaacagataacggcgtaaataag -3'

nucS3

Bgl II Eco RV

5'-cgagatctgatatccatcacaaacagataacggcgtaaataag -3'

nucR

Bam HI

5'-cgggatccttatggacctgaatcagcgttgctc -3'

NucSeq

5'-ggatgctttgtttcaggtgtatc -3'

pTREP_F

5'-catgatatcggtacctcaagctcatatcattgtccggcaatgggtgtgggctttttttgttttagcggataa
caatttcacac -3'

pTREP_R

5'-gcggatcccccggttaattaatgttttaaacactagtcgaagatctcgcaattctcctgtgtgaaatt
gttatccgcta -3'

pUC_F

5'-cgccagggttttccagtcacgac -3'

V_R

5'-tcaggggggaggagcctatg -3'

V₁

5'-tcgtatgttggtggaattgtg -3'

V₂

5'-tccggctcgtatgttggtggaattg -3'

FIG. 3

pTREP-Nuc vectors allow cloning of genomic DNA into each frame with respect to the nuclease gene

(i)

pTREP1-nuc1 (EcoRV)	AAGTATCAGATCT-- <u>GATATC</u> --TCACAAACAGATAACGGCGTAAAT	Frame=+1
▲.....	
pTREP1-nuc2 (Sma I)	AAGTATCAGATCTT <u>CCCCGGA</u> --TCACAAACAGATAACGGCGTAAAT	Frame=+2
▲.....	
pTREP1-nuc3 (EcoRV)	AAGTATCAGATCT-- <u>GATATC</u> TCACAAACAGATAACGGCGTAAAT	Frame=+3
▲.....	
Nuclease Gene	TCACAAACAGATAACGGCGTAAAT	

Cloning site is indicated by an arrow

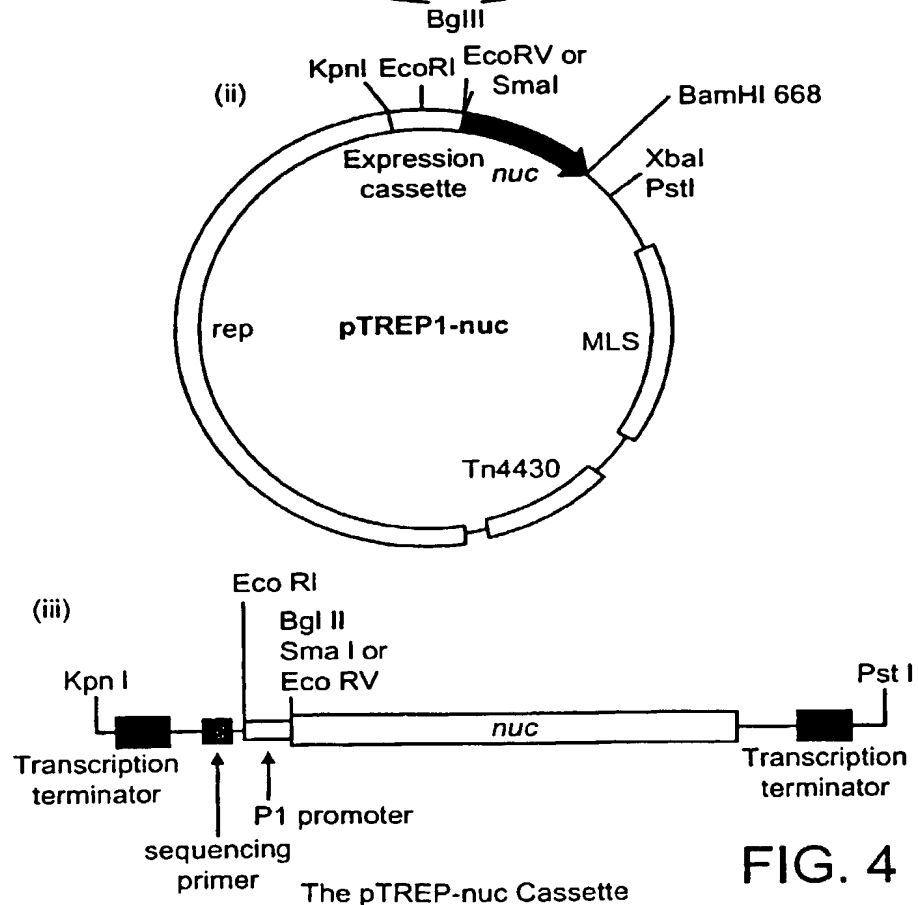
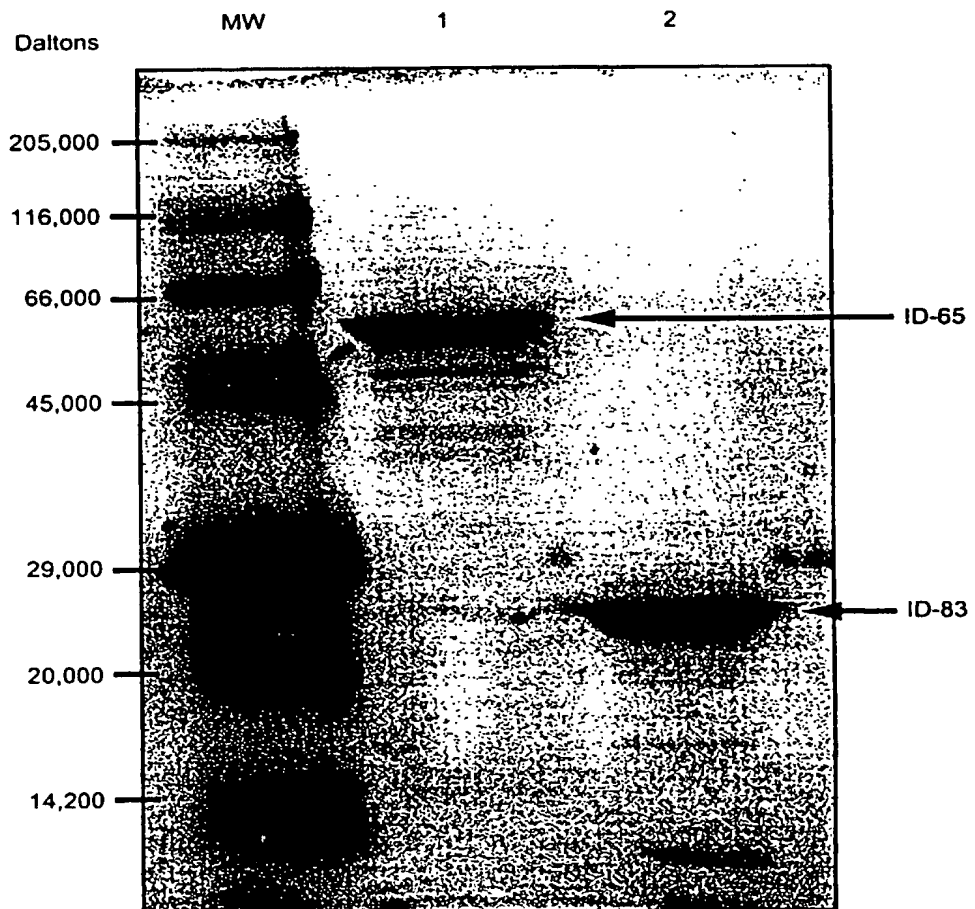


FIG. 4

102 / 110

FIG. 5

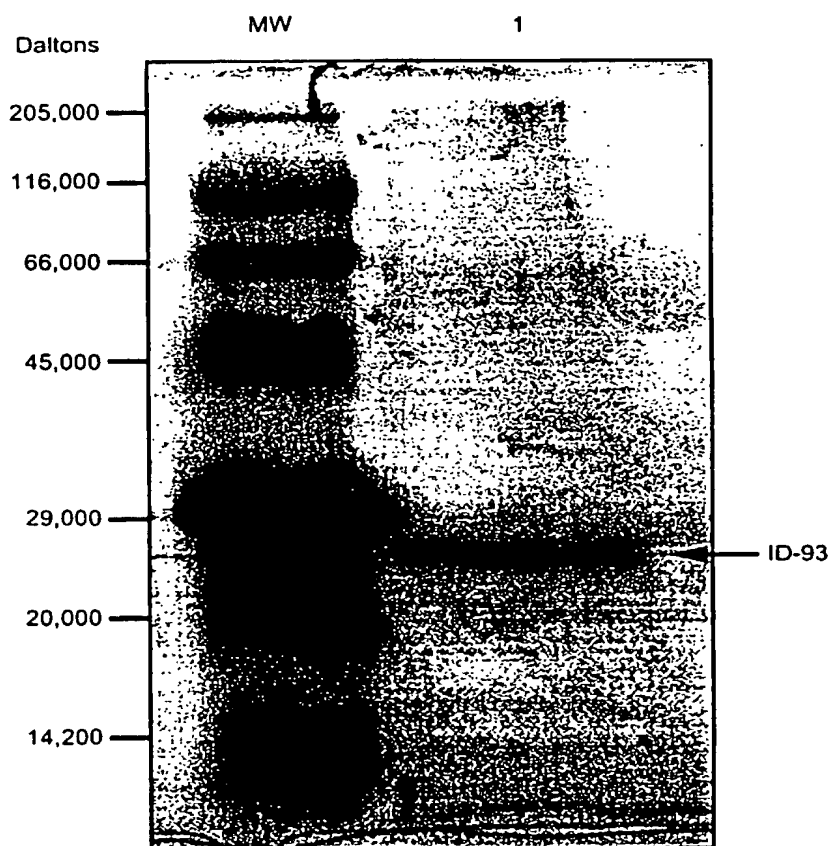
SDS-PAGE analysis of the purified ID-65 and ID-83 protein antigens



103 / 110

FIG. 6

SDS-PAGE analysis of the purified ID-93 antigen



104 / 110

FIG. 7

SDS-PAGE analysis of the purified ID-89 and ID-96 protein antigens

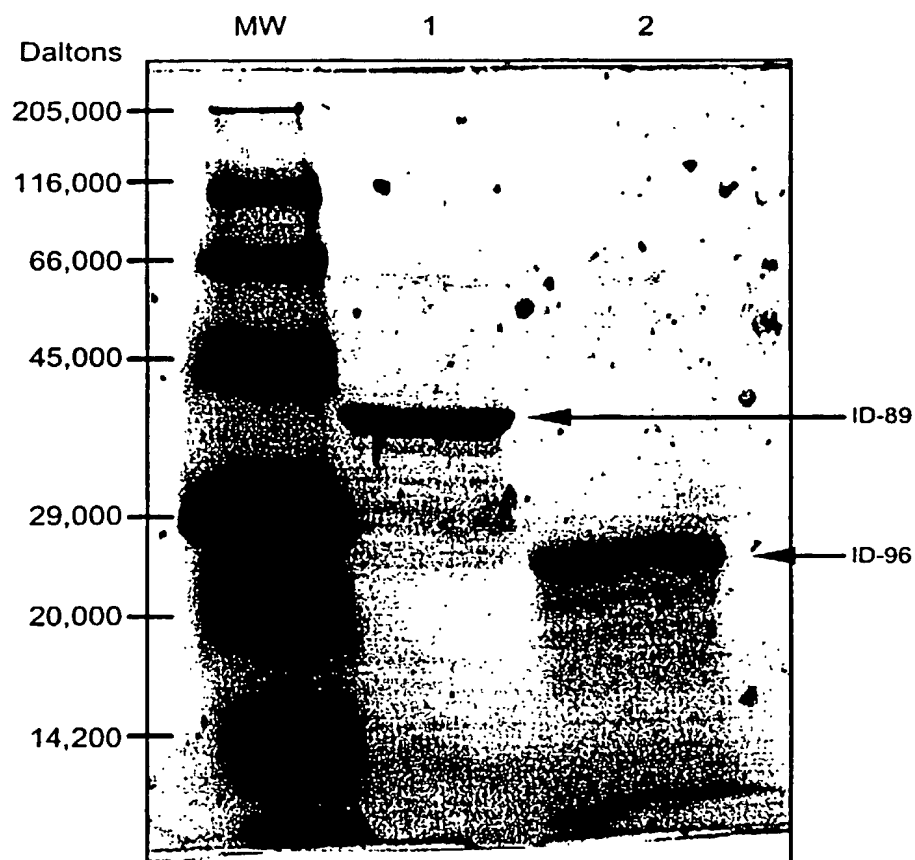
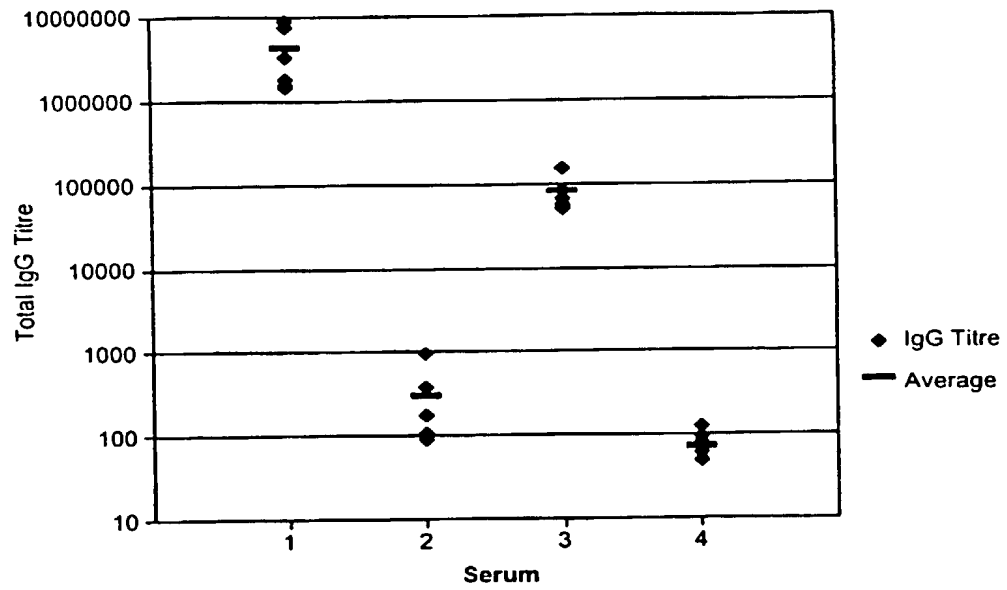


FIG. 8

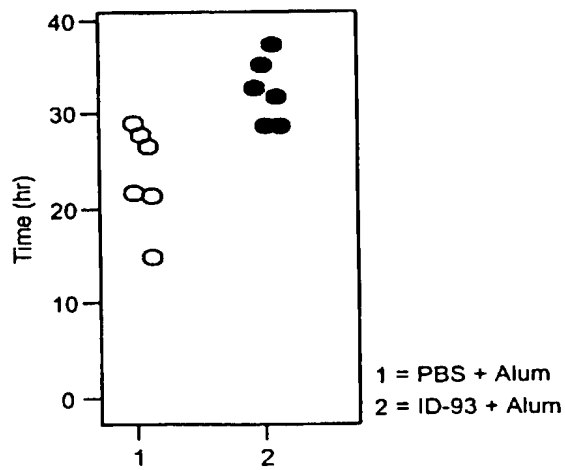
IgG Titres against the ID-65 and ID-83 proteins

ID-65 and ID-83 Vaccinations -IgG Titres

**FIG. 9**

Survival data

ID-93 Vaccination- GBS Challenge and Survival



106 / 110

FIG. 10

IgG Titres against the ID-93 protein

ID-93 Protein Vaccine -IgG Titres

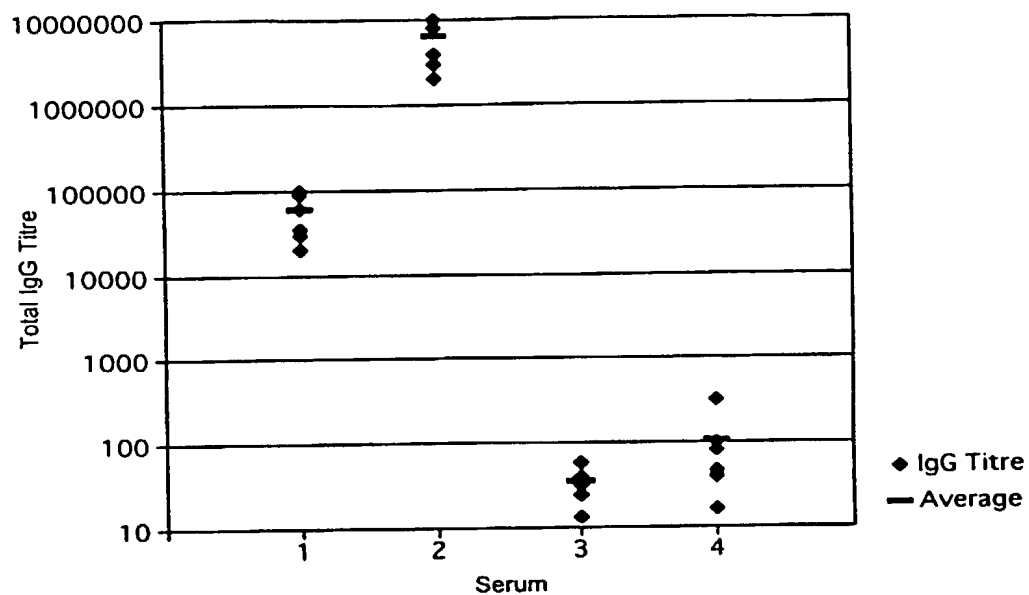
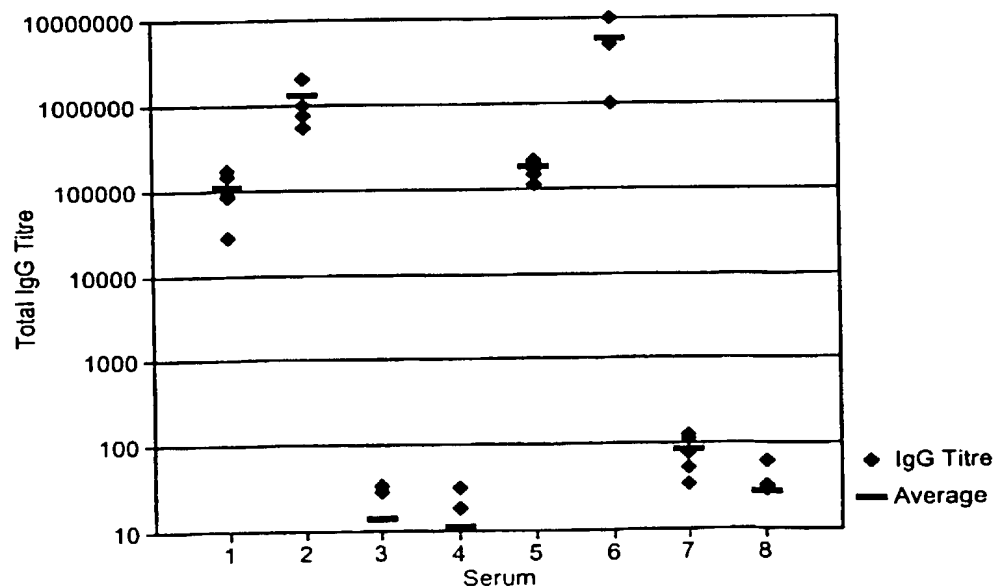
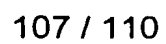


FIG. 11

IgG Titres against the ID-89 and ID-96 proteins

ID-89 and ID-96 Protein Vaccines -IgG Titres





108 / 110

FIG. 13

Southern blot analysis - ID-65

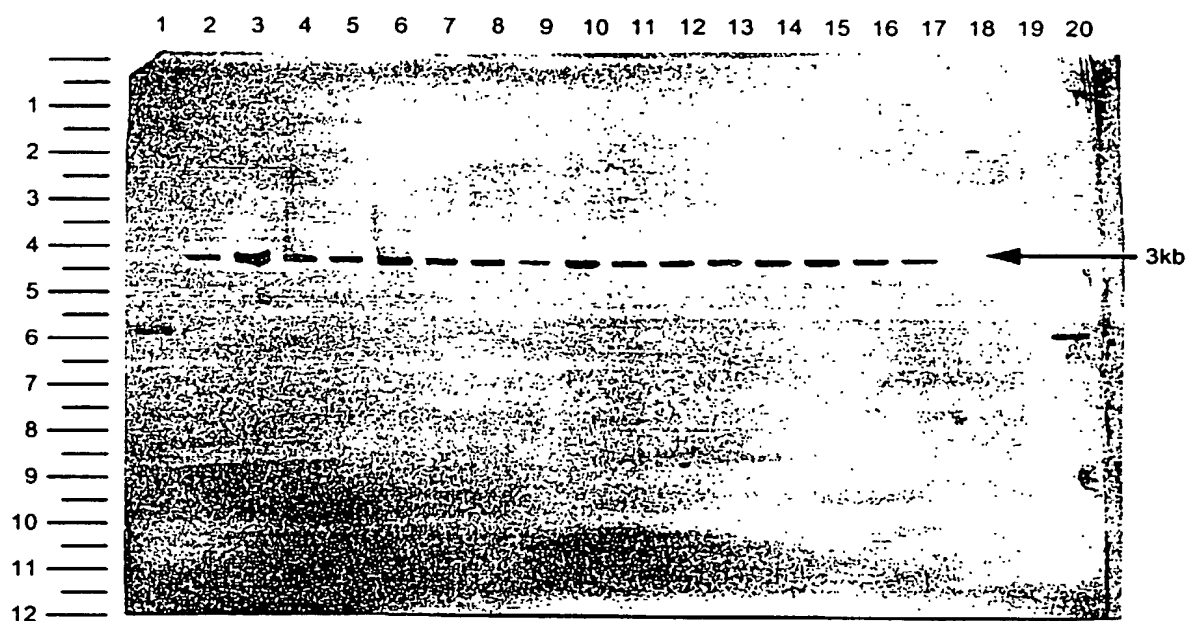


FIG. 14

Southern blot analysis - ID-89

